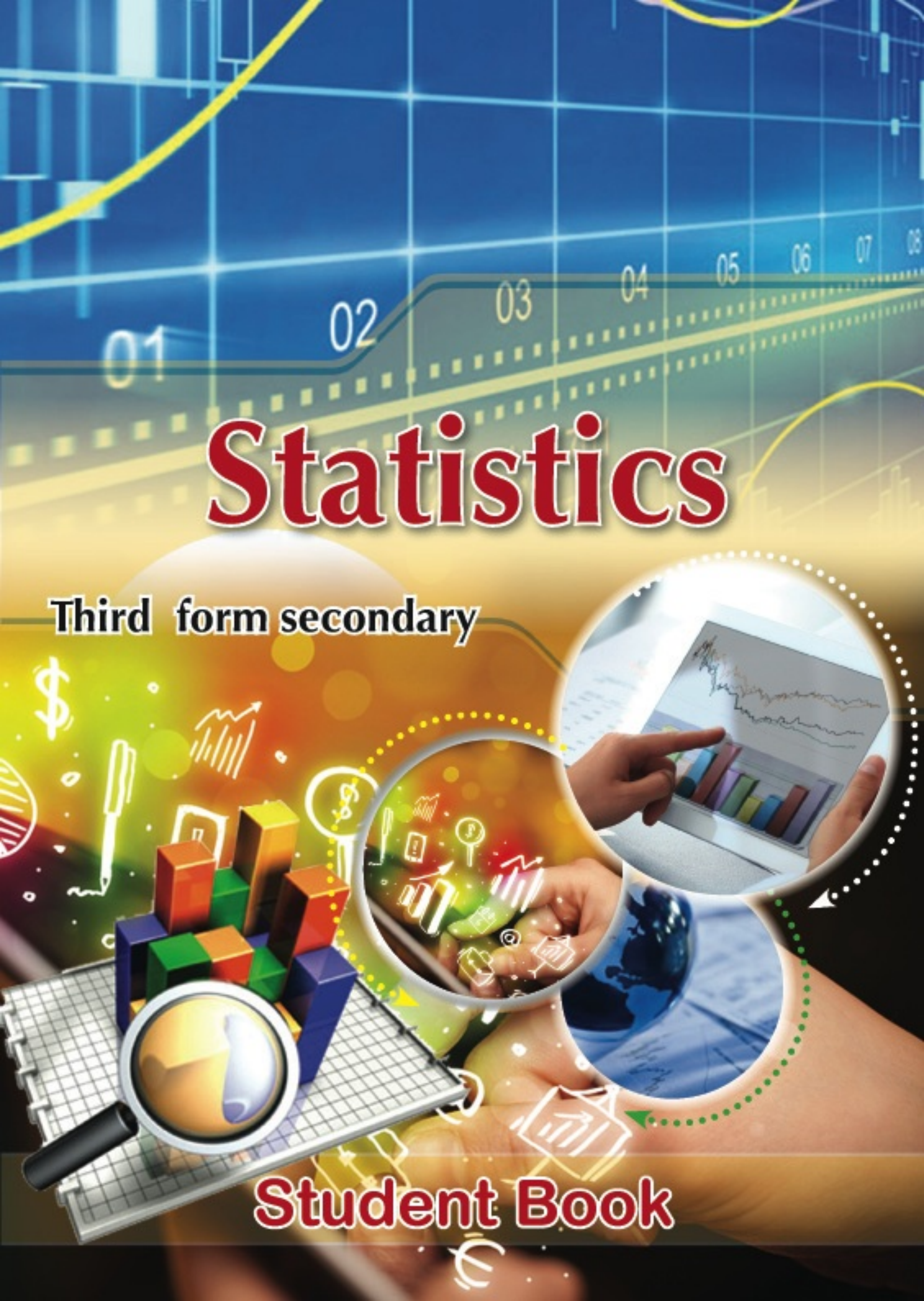
Statistics

Third form secondary



Student Book

Statistics

Third form secondary

Student Book

Name : ....

Class : ....

School: ....

Authors

Mr. Kamal Younis Kabsha

Prof.Dr. Ahmed Kamal El-kholy

First eddition 2016 / 2017

Deposit number 8701 / 2016

I.S.B.N 5 - 029 - 706 - 977 - 978

SAKKARA PUBLISHING

Introduction

بسم الله الرحمن الرحيم

يسعدنا ونحن نقدم هذا الكتاب أن نوضح الفلسفة التى تم فى ضوئها بناء المادة التعليمية ونوجزها فيمايلى:

١ تنمية وحدة المعرفة وتكاملها فى الرياضيات، ودمج المفاهيم والترابط بين كل مجالات الرياضيات المدرسية.

٢ تزويد المتعلم بما هو وظيفى من معلومات ومفاهيم وخطط لحل المشكلات.

٣ تبنّى مدخل المعايير القومية للتعليم فى مصر والمستويات التعليمية وذلك من خلال:

أ) تحديد ما ينبغى على المتعلم أن يتعلمه ولماذا يتعلمه.

ب) تحديد مخرجات التعلم بدقة، وقد ركزت على مايلى:

أن يظل تعلم الرياضيات هدف يسعى المتعلم لتحقيقه طوال حياته - أن يكون المتعلم محبًّا للرياضيات ومبادرًا بدراستها - أن يكون المتعلم قادرًا على العمل منفردًا أو ضمن فريق - أن يكون المتعلم نشطًا ومثابرًا ومواظبًا ومبتكرًا - أن يكون المتعلم قادرًا على التواصل بلغة الرياضيات.

٤ اقتراح أساليب وطرق للتدريس وذلك من خلال كتاب (دليل المعلم).

٥ اقتراح أنشطة متنوعة تتناسب مع المحتوى ليختار المتعلم النشاط الملائم له.

٦ احترام الرياضيات واحترام المساهمات الإنسانية منها على مستوى العالم والأمة والوطن، وتعرف مساهمات وإنجازات العلماء المسلمين والعرب والأجانب.

وأخيرًا ..نتمنى أن نكون قد وفقنا فى إنجاز هذا العمل لما فيه خير لأولادنا، ولمصرنا العزيزة.

والله من وراء القصد، وهو يهدى إلى سواء السبيل

Contents

Unit one: Correlation and Regression

1 - 1 Correlation 4

1 - 2 Regression 16

Unit summary 26

General exercises 27

Accumulative test 29

Unit two: Conditional Probabilit

2 - 1 Conditional Probability 32

2 - 2 Independent Events 41

Unit summary 48

General exercises 49

Accumulative test 51

Contents

Unit three: Random Variables and Probability Distributions

3 - 1 Discrete random variable 54

3 - 2 Expectation and Variance of a Discrete Random Variable 61

3 - 3 Probability density function of the continuous random variable 68

Unit summary 73

General exercises 75

Accumulative test 77

Unit four: Normal distribution

4 - 1 Normal distribution 80

4 - 2 Some practical applications of the normal distribution 94

Unit summary 100

General exercises 101

Accumulative test 103

General tests 105

Unit 1

Correlation and Regression

introduction

Statistics is an important branch of mathematics which has various applications since it concerns with gathering, representing and reducing the data in the form of digital indicators to describe, measure and analyze the basic features of such data in order to take the right decisions. Statistics has wide applications in all fields of physical, human, economic and social sciences. In this unit, we are going to focus on analyzing the bivariate data and studying the degree of the direction of the relation between the two variables and the shape of this relation. Initially, we study the correlation which reveals the degree and the power of the relation between the two variables. This relation may be in the form of a direct or inverse form. It is worth mentioning that the correlation studies the relation and its direction between a variable and another. It is necessary to realize that the relation does not refer to causation because it does not refer to the existence of an effect to a variable on the other as it will be shown in Lesson 1 in this Unit. In this unit also, you are going to study the simple linear regression which concerns with estimating the form of this relation through which you can predict the value of the dependent variable if the value of the independent variable is known. It gets more accurate, the more the sample randomly chosen. You will study some modern technologies about the scientific calculators and the computer statistical programs such as(SPSS) to perform the calculations and to graph the data related to the correlation and the linear regression between two phenomena.

Unit objectives

By the end of the unit and carrying out the involved activities, the students should be able to

- Identify what is meant by correlation between two variables.
- Calculate the correlation coefficient between two variables in different ways (Pearson's method and Spearman's method) and interpret them mathematically.
- Understand the meaning of the regression line and estimate its importance to study the relation between two variables.
- Represent the relation between two variables in a Cartesian diagram and judge from the graph the existence of the relation and the power of its degree.
- Identify the meaning of the liner regression coefficient and interpret what can be deduce by knowing the value of such a coefficient.
- Find the regression line equation of any of the two variables on the other using the least squares method.
- Use the calculator and computer to perform the mathematical calculations and graph the data of the correlation and the linear regression between two phenomena.
- Use a given regression line equation to predict the value of one of the two variables in terms of the corresponding value of the other variable.
- Apply the correlation and the linear regression in research situations.
- Appreciate contributions of using the correlation and the linear regression in solving social and daily life problems.

Key Terms

- *Correlation*
- *Regression*
- *Linear correlation*
- *Correlation Coefficient*
- *Direct Correlation*
- *Inverse Correlation*
- *Scatter diagram*
- *Pearson Correlation coefficient*
- *Spearman Correlation coefficient*
- *Regression Line*
- *Least Square*

Lessons of the unit

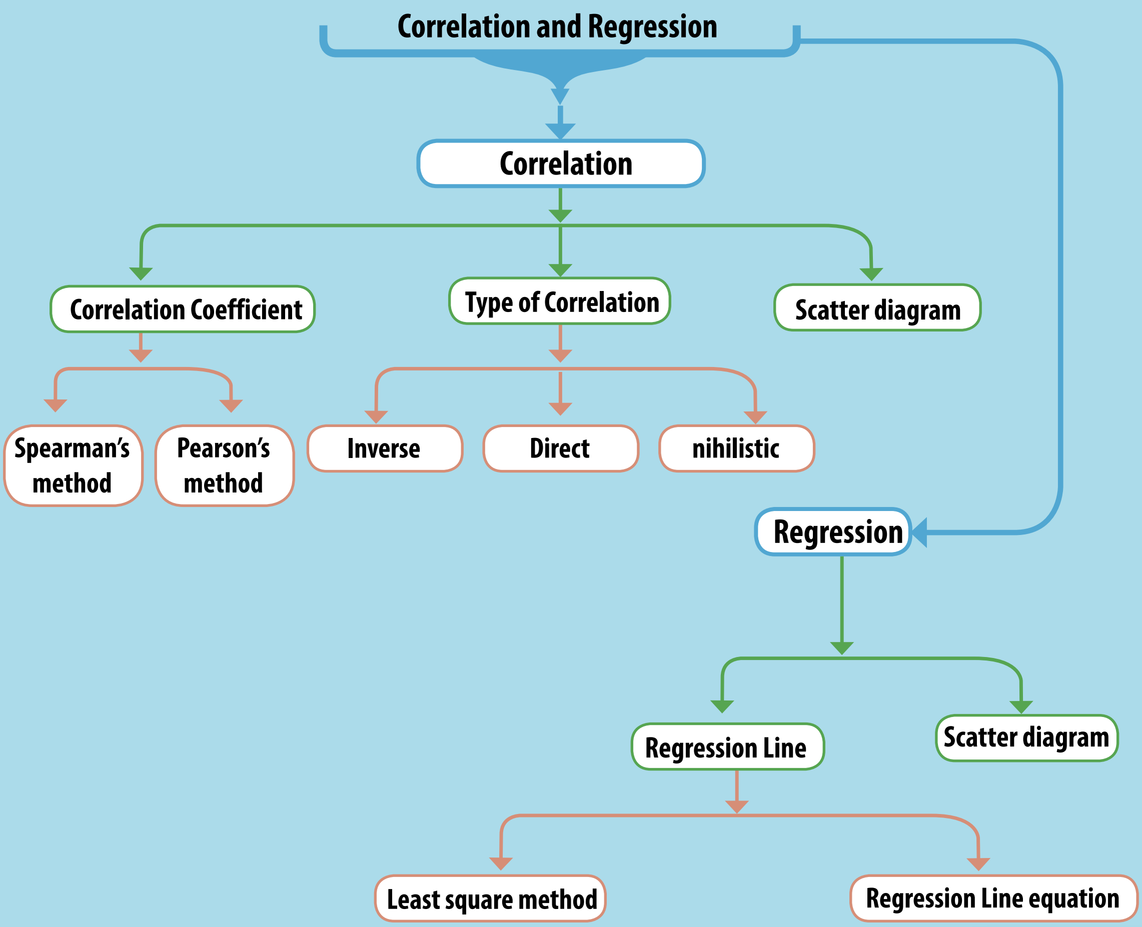
Lesson (1-1) Correlation

Lesson (1-2) Regression

Materials

Scientific calculator - excel program - Spss program

Chart of the unit



Unit one

1 - 1 Correlation

You will learn

- Definition of correlation
- Scatter diagram
- Direct correlation and inverse correlation
- Linear correlation coefficient
- Pearson's linear correlation coefficient
- Spearman's rank correlation coefficient

Key terms

- Correlation
- Linear Correlation
- Correlation Coefficient
- Direct Correlation
- Inverse Correlation
- Scatter diagram
- Pearson Correlation Coefficient
- spearman`s rank correltion coefficient

Introduction:

You have previously learned in Statistics how to describe a set of data representing a phenomenon using some statistical scales such as the scales of the central tendency, dispersion and the coefficient of variation. In this lesson, you are going to learn how to describe the items of two different phenomena in accordance to the relation between both. In other words, if one of the two variables has changed in a certain direction(by increasing or decreasing), the other variable leans to change in a certain direction by increasing or decreasing, too. In this case, the correlation is called a direct correlation. If one of the two variable changes in the direction of increasing and the other changes in the direction of decreasing and vice versa, the correlation in this case is called the inverse correlation.

Correlation:

Think and discuss

meditate the following examples and record your observations:

1- The relation between the side length and the area of the square.

2- The relation between the blood pressure and the age.

3- The increase in the price of an item and the demands of buying it.

4- The decrease of temperature and the demand to consume the fuel.

5- The relation between rising of the sea level and the rise of the temperature.

From the examples above, we observe that:

- The two correlated variables change in the same direction. I.e. the increasing or decreasing of a variable leads to the increasing or decreasing of the other as in examples 1,2 and 3. It is said the correlation between both variables is positive (direct).

materials

- Scientific calculator

In examples 4 and 5, we notice that two correlated variables inversely change. In other words, the increasing or decreasing of one variable leads to decreasing or increasing of the other variable. hence, it is said that the correlation between them is negative (inverse).

Definition Correlation is a statistical method by which the degree and the type between two variables can be determined.

The relation between the two variables ranges from a strong degree to a weak one. When the relation between the two variables is strong, it means that knowing the value of one variable will help to predict the value of the other variable. On the contrary, when the relation is weak, it means that knowing the value of one variable will not help to predict the value of the other variable.
One of the most important methods helping us identify the degree and type of the relation between the two variables is the Scatter Diagram.

Scatter diagram

Definition Scatter Diagram is a graphical representation of a number of the ordered pairs (x,y) to describe the relation between two variables.

If we denote the first phenomenon by the symbol (x) and the second phenomenon by the symbol (y), then the next diagrams illustrate the relation between x and y which show the scatter diagram.

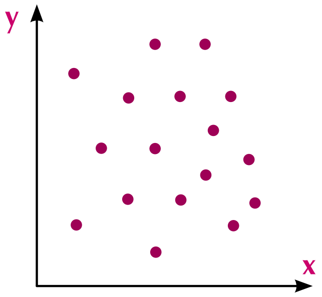


Figure 1
There is no correlation

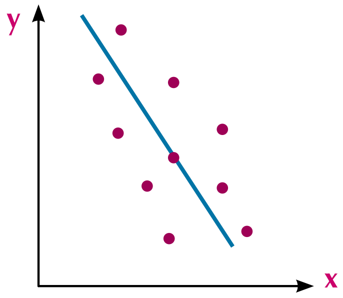


Figure 2
There is an inverse linear correlation

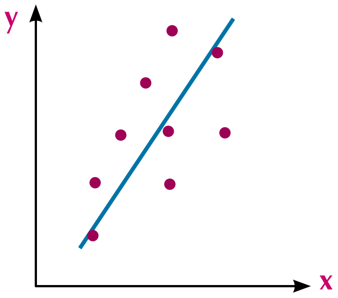


Figure 3
There is a direct linear correlation

Linear Correlation

Definition The simple linear correlation is known as the measure of the degree of the relation between two variables.

Activity

graph the scatter diagram for each of the following data mentioning the type of the relation expressing such data.

1.

x	7	8	9	10	11	12
y	13	14	17	18	21	23

2.

x	3	4	7	8	11	15
y	23	22	20	18	17	16

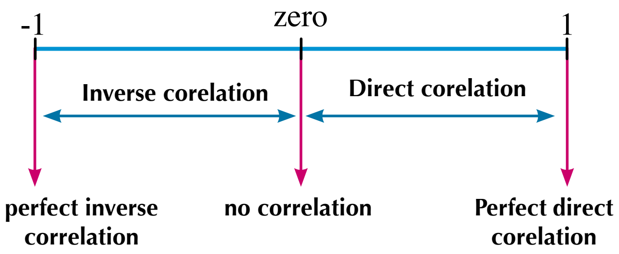
3						
x	7	9	11	13	15	16
y	14	7	20	6	12	10

Correlation Coefficient

The correlation coefficient is denoted by the symbol (r). It is a relatively quantitative scale measuring how much the two variables are correlated where $-1 \leqslant r \leqslant 1$. It is said the correlation is perfect direct if the correlation coefficient r = 1, is perfect inverse if the correlation coefficient r = -1 and the correlation is nihilistic when r = 0.

We notice that:

The nearer the value of the correlation coefficient to 1 is, the stronger the direct correlation between the two variable is and the nearer its value to zero is, the weaker the direct correlation is and we will say the same in the inverse correlation. The opposite figure shows that.



Oral expression: Multi -choice:

The strongest correlation coefficient of the following is:

a – 0.8 b – 0.5 c 0.4 d 0.7

Pearson`s Correlation coefficient

Suppose we have a set of (n) people and we get data from those people about the values of two variables x and y. The data we get ought to be on the form:

The values of first variable x: $x_1 , x_2 , x_3 , \ldots\ldots\ldots\ldots , x_n$

The values of second variable y: $y_1 , y_2 , y_3 , \ldots\ldots\ldots\ldots , y_n$

If we denote the correlation coefficient by the symbol (r), then Pearson's correlation coefficient between the two variables x and y or the linear correlation coefficient can be found from the relation:

$$r = \frac{n\,\Sigma x\,y - (\Sigma x \times \Sigma y)}{\sqrt{n\,\Sigma x^2 - (\Sigma x)^2}\,\sqrt{n\,\Sigma y^2 - (\Sigma y)^2}}$$

Where: "Σ" is the symbol of summation notation and read as the sum.

n denotes the number of the items.

$\Sigma x = x_1 + x_2 + x_3 \ldots\ldots + x_n$,

$\Sigma y = y_1 + y_2 + y_3 \ldots\ldots + y_n$,

$\Sigma x\,y = x_1 y_1 + x_2 y_2 + x_3 y_3 \ldots\ldots + x_n y_n$

$\Sigma x^2 = x_1^{\,2} + x_2^{\,2} + x_3^{\,2} \ldots\ldots + x^2_n$,

$\Sigma y^2 = y_1^{\,2} + y_2^{\,2} + y_3^{\,2} \ldots\ldots + y_n^2$

Example

1 The next table shows the marks of 10 students in the history and geography subjects.

history x	75	80	93	65	87	71	98	69	84	78
geography y	82	78	86	72	91	80	95	73	89	74

Required is to calculate the Pearson's correlation coefficient between the two variables x and y and identify the type of the correlation.

Solution

We form the next table:

x	y	x^2	y^2	x y
75	82	5625	6724	6150
80	78	6400	6084	6240
93	86	8649	7396	7998
65	72	4225	5184	4680
87	91	7569	8281	7917
71	80	5041	6400	5680
98	95	9604	9025	9310
69	73	4761	5329	5037
84	89	7056	7921	7476
78	74	6084	5476	5772
$\Sigma x = 800$	$\Sigma y = 820$	$\Sigma x^2 = 65014$	$\Sigma y^2 = 67820$	$\Sigma xy = 66260$

$$\because r = \frac{n\,\Sigma x y - (\Sigma x \times \Sigma y)}{\sqrt{n\,\Sigma x^2 - (\Sigma x)^2}\ \sqrt{n\,\Sigma y^2 - (\Sigma y)^2}}$$

$$\therefore\ r = \frac{10 \times 66260 - (800 \times 820)}{\sqrt{10 \times 65014 - (800)^2}\sqrt{10 \times 67820 - (820)^2}}$$

$$= \frac{6006}{\sqrt{10140}\ \sqrt{5800}} \simeq 0.8606$$

The correlation is direct .

Try to solve

1 **From the data of the next table:**

x	20	23	24	25	28	30
y	35	31	30	27	29	28

Calculate the Pearson's correlation coefficient (linear) between the two variables x and y and identify its type

Using the scientific calculator:

A lot of scientific calculators in the markets are supporting to find the sums of the columns in the table above and calculate the correlation coefficient as follows:

- **Preparing the calculator for the statistics system:**

  By pressing MODE then 3

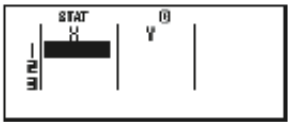
Statistical and regression calculations	MODE 3 (STAT)

  Select from the drop-down list:

Paired-variable (X, Y), linear regression (y = A + Bx)	2 (A+BX)

- **Entering the data:**

  MODE 3 (STAT) 2 (A+BX)

  Fill the shown table in the figure in all the values of (x and y) by typing the number existed in a table After typing the number, press = until you finish typing all the values of (x and y).

- **Recalling the products:**

  Press the buttons. SHIFT 1 (STAT) to get 3:sum

  And select from this list each of n:

  1 : Σx^2 , 2 : Σx , 3 : Σy^2 , 4 : Σy , 5 : Σxy

  By pressing the buttons from 1 to 5 severally liable.

To find the correlation coefficient (r) press the following buttons:

(STAT), from the drop-down list, press 5 : Reg

And from the drop-down list, press 3: r to get the value of the correlation coefficient required between the two variables x and y.

Activity

Use the calculator to check the answer of the example above

The program (SPSS)

The program (SPSS) is the abbreviation of (Statistical Package for Social Science). It is a set of comprehensive mathematical packs or data to analyze such data. This program is used in the scientific researches involving digital data.

The program can read all the data of different files, analyzing them, getting the results and giving statistical reports. The programs also allows the users to edit the data in the form of variables and new data using the equation. You can also use the program to save the data in files and naming or editing the names of the data files. You can also pack up the data and files through controlling the order and selection list available in this program. This can be done to include all the stages of analyzing the data and the statistical process via four main steps:

1 - Coding the data.
2 - Placing the data in the program.
3 - Selecting the proper form and testing and analyzing the data.
4 - Identifying the variable data to be analyzed and to fulfill the statistical process.

Operating the SPSS program:

You can start the SPSS program by pressing the window (START) in the main list, then go to the (program) list and search for the SPSS program within two times to open the program.

Components of the programs and their function:

Command list:

It is a special task bar for this program where the user can select the order which he /she wants through pressing the icon of each statistical order and then the results are viewed in the report board. The command list contains nine main orders. When you press any of them, they get branched into sub- orders besides, the (help) icon.

Data view:

It is a screen where the user can control adding or cancelling the dependent data for each variable. The user can add any independent variable in a column on the data view where the user can convert to view or watch the variables by pressing and moving between the two orders (Variable View) and (Data View) placed on the left of the variable view.

Variable view:

The view of defining the variable data which contains parallel columns and each column contains special data for each variable. To view the definition of each variable, the user ought to use the mouse to do (Double Click) or to press the order (Variable View) placed under the definition view. At this very minute, the shape of the view get changed and the definition bar appears:

Name	Type
Width	Values

When you press the definition bar, the coding (denotation)appears. After that, press the button (Add) to view the value of the symbol and position.

Steps you can control:

(1) **The ability to recall the previous data:** The user can control the process of recalling the data and files by pressing the button (File), then the button (Open). The user selects the file containing the data wanted to be recalled which contain the statistical reports done before and finally presses the button (Save).

(2) **Saving the new changes in a file:** The user can save the new changes in a file by pressing the order (Save) or the order (Save as) to save and name the new selected file.

(3) Adding edits and managing the variables: The user goes to the (Data Editor) and add the data he /she wants to be able to:

- edit the value of the data.
- Defining the variables; identifying the type of the added data, the economic indicators and all the variables.

(4) The user can add a new variable view and watch the arrange of the changes happened by using the main order (Data), then follow each variable that the user wants such as adding a variable or adding a new view or editing the arrangement of the data.

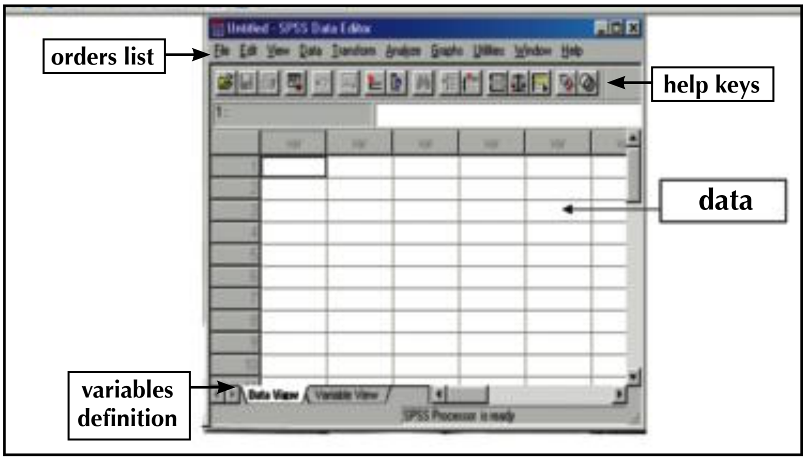
(5) Forming a fully new variable using an equation where the user goes to (Transform), then moving to (Compute) and identifies the name of the new variable in the list (Targer Variable).

(6) The ability to cancel any variable or view.

(7) Ordering the view where the program sets up a new variable containing a serial number in order to order the views ascendingly or descendingly.

(8) Performing a statistical process identifying and graduating the statistical description and the repetition of data.

(9) The ability of graphing the variables through setting up a graph to view the variable analysis and interpret what happened to the new variables.



Activity

You can visit the next site to download the SPSS: *http://www-01.ibm.com/software/analytics/spss/* to check the example above.

Example

(2) Find the pearson's correlation coefficient between the two variables x and y and identify its type if.

$\Sigma x = 68$ $\Sigma y = 36$ $\Sigma xy = 348$

$\Sigma x^2 = 620$ $\Sigma y^2 = 204$ $n = 8$

Solution

$$\because r = \frac{n\,\Sigma xy - (\Sigma x \times \Sigma y)}{\sqrt{n\,\Sigma x^2 - (\Sigma x)^2}\,\sqrt{n\,\Sigma y^2 - (\Sigma y)^2}}$$

$$\therefore r = \frac{8 \times 348 - (68 \times 36)}{\sqrt{8 \times 620 - (68)^2}\,\sqrt{8 \times 204 - (36)^2}} = \frac{336}{\sqrt{336}\,\sqrt{336}} = 1$$

The value of the correlation coefficient (+1) means that the relation between the two variables x and y is perfect direct.

Try to solve

2) Find the pearson's correlation coefficient between the two variables x and y and identify its type if.

$\Sigma x = 92$ $\qquad \Sigma y = 36$ $\qquad \Sigma xy = 372$

$\Sigma x2 = 1100$ $\qquad \Sigma y2 = 204$ $\qquad n = 8$

Spearman's Rank Correltion Coefficient

Think and discuss

A statistician has studied the relation among the degrees of two school subjects for seven students and has recorded the results in the next table:

Subjects 1	Weak	Pass	Weak	good	Weak	excellent	Very good
Subjects 2	Weak	Pass	good	Pass	Weak	Very good	Pass

can you help the statistician , to realize the relation between those two subjects and find the correlation coefficient between them?

We cannot use Pearson's correlation coefficient in "Think and discuss" since it depends on the quantitative (numerical)data only. In this case, the data are descriptive (as in the previous title). another correlation coefficient can be used _ Spearman's rank correlation coefficient. It gives a measure to the correlation in both the descriptive and quantitative data which are ordered as in the previous title. This coefficient relies on the order of the variable values taking into consideration the ascending and descending order. Then, we use the next relation:

$$r = 1 - \frac{6\Sigma D^2}{n(n^2 - 1)}$$

Where D is the difference between the order of the two variables x and y and n is the number of the values of each variable.

Notice that:

- Spearman's correlation coefficient can be calculated whether the data are quantitative or descriptive, whereas Pearson's correlation coefficient can be calculated only if the variables are quantitative.
- Spearman's rank correlation coefficient is characterized with its easiness even the data are not ordered.
- The Spearman's correlation coefficient does not regard the difference of the numbers when the ranks are calculated and though, it is less accurate.

Example

3 Find Spearman's rank correlation coefficient in the previous "Think and discuss" and identify its type.

Solution

In this example, we order the two phenomena a regular ascending order by giving each student a rank for a subject so does the other subject for the same student as shown in the next table.

Subjects 1	Weak	Pass	Weak	good	Weak	excellent	Very good
Ordering with repetition	1	4	2	5	3	7	6
Final repetition	2	4	2	5	2	7	6

We notice that the case "weak" has repeated three times and occupied the places 1,2 and 3.

hence, the rank of each = $\frac{1+2+3}{3}$ = 2 (it is the arithmetic mean of the numbers 1,2 and 3) , and similarly,

Subjects 2	Weak	Pass	good	Pass	Weak	Very good	Pass
Ordering with repetition	1	3	6	4	2	7	5
Final repetition	1.5	4	6	4	1.5	7	4

We notice that the level "weak" has repeated twice and occupied the places 1 and 2.

hence, the rank of each = $\frac{1+2}{2}$ = 1.5 it is the arithmetic mean of the numbers 1 and 2).

Also, the level"pass" has repeated three times and occupied the places 3,4 and 5.

hence, the rank of each = $\frac{3+4+5}{3}$ = 4 We summarize the solution in the table as follows:

x	y	Ranks of x	Ranks of y	D	D^2
Weak	Weak	2	1.5	0.5	0.25
Pass	Pass	4	4	zero	zero
Weak	good	2	6	– 4	16
good	Pass	5	4	1	1
Weak	Weak	2	1.5	0.5	0.25
excellent	Very good	7	7	zero	zero
Very good	Pass	6	4	2	4
					21.5

$\because$ r $= 1 - \frac{6\Sigma D^2}{n(n^2-1)}$ $\because$ $r = 1 - \frac{6 \times 21{,}5}{7\ (49 - 1)}$

$= 1 - \frac{129}{336} \simeq 0{,}6161$ **It is a direct correlation.**

Try to solve

3. In a study about the relation between the students' levels in statistics and mathematics, the degrees of six students have been as follows:

Degrees of statistics(x)	Pass	Very good	excellent	Very good	Pass	Pass
Degrees of mathematics (y)	good	good	Very good	excellent	good	Weak

calculate the Spearman's rank correlation coefficient among the degrees and identify its type.

Example

4. calculate the Spearman's rank correlation coefficient between x and y through the data of the next table:

x	4	7	8	5	8	12
y	7	6	6	4	6	10

Solution

Form the next table:

x	y	ranks x	ranks y	D	D^2
4	7	6	2	4	16
7	6	4	4	0	0
8	6	2.5	4	–1.5	2.25
5	4	5	6	–1	1
8	6	2.5	4	–1.5	2.25
12	10	1	1	0	0
					21.5

$\because$ $r = 1 - \frac{6\Sigma D^2}{n(n^2-1)}$ $\therefore 1 = \frac{6 \times 21{,}5}{6\ (36 - 1)}$ $\simeq 0{,}3857$ The correlation is direct

Critical thinking: Does ΣD^2 vary if we order the two phenomena x and y ascendingly? explain your answer

Try to solve

4. Calculate the Spearman's rank correlation coefficient between x and y and identify its type from the data shown in the next table:

x	10	7	8	7	6	4
y	5	8	7	9	9	10

Exercises 1 - 1

First: Choose the correct answer:

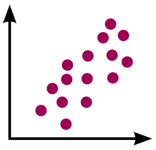
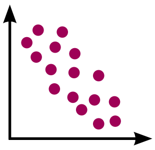
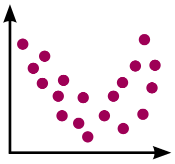
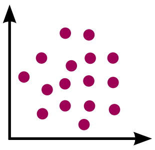
1. The strongest correlation coefficient of the following is:

   a) – 0.94 b) zero c) 0.5 d) 0.85

2. The strongest inverse correlation coefficient of the following is:

   a) – 0.2 b) – 0.5 c) – 0.7 d) – 0.8

3. the scatter diagram representing inverse correlation is:

   a)  b)  c)  d) 

4. The weakest correlation coefficient of the following is:

   a) - 1.2 b) - 0.7 c) 0.12 d) 0.9

5. One of the next numbers can represent the strongest inverse correlation coefficient between two variables:

   a) 0.3 b) 0.9 c) - 1.1 d) - 0.95

Answer the following questions

6. From the data of the next table

x	12	10	14	11	12	9
y	18	17	23	19	20	15

First: Calculate Spearman's rank correlation coefficient between the two variables x and y.

Second: Calculate Pearson's linear correlation coefficient between x and y.

7. From the data of the next table

x	7	7	8	3	7	11
y	8	4	12	2	10	11

Calculate Spearman's rank correlation coefficient between the two variables x and y.

8. From the data of the next table

x	1	3	4	6	7	9
y	6	4	4	3	2	1

Calculate Pearson's correlation coefficient between the values of x and y and show its type.

9. From the data of the next table

x	6	5	7	8	10	6	7
y	4	7	5	6	8	7	8

Calculate Pearson's correlation coefficient between the values of x and y and show its type.

10 From the data of the next table

x	3	1	6	4	3	8
y	7	4	5	8	6	7

Calculate Spearman's rank correlation coefficient between x and y and show its type.

11 From the data of the next table

x	Very good	Very good	good	Weak	Pass	Very good
y	good	Pass	good	excellent	Very good	Pass

Calculate Spearman's rank correlation coefficient between x and y.

12 Find Pearson's correlation coefficient between the values of x and y and show its type if:

$\Sigma x = 220$ $\Sigma y = 140$ $\Sigma xy = 2658$

$\Sigma x^2 = 5486$ $\Sigma y^2 = 2292$ $n = 10$

13 Trade: The following table shows a 6 - book group in regard to its price (x) and the sales (y).

Price (x)	Low	Very low	Average	Very high	high	Very high
Sales (y)	high	high	Very high	Low	Average	Low

Calculate Spearman's rank correlation coefficient between the price and sale of the book.

14 Advertisement: An advertising company has decided to study the relation between its expenditure on advertisement x (in thousand L.E) and its sales y (in thousand units). If the data of the company's eight branches are as follows:

x	19	18	7	10	4	13	15	5
y	12	10	7	9	6	13	14	12

Find the rank correlation coefficient between the expenditure and sales and show its type.

15 Education The next data represent the marks of ten students in chemistry and biology.

chemistry	60	85	55	90	65	50	80	70	95	75
biology	55	75	50	95	60	65	85	80	90	70

Calculate Pearson's linear correlation coefficient and determine its type.

16 Birth: In a study to determine the relation between the mother's age and the number of her kids, the following data have been determined:

mother's age	18	20	23	27	29	32	33	35
number of kids	2	1	1	2	3	4	3	5

Calculate spearman's rank correlation relation and determine its type.

Unit one
1 - 2 Regression

Regression

You will learn

- Definition of regression
- Types of regression
- Regression line equation
- Method of least squares
- Activities on finding regression line equation

Key terms

- Regression
- Regression line
- Least squares

Introduction:

You have previously studied the function and identified its graph. In the previous lesson, you learned the scatter diagram and you also realized that the objective of graphing it is to identify the nature of the relation between two variables x and y through the data related to both variables. You also learned that the properties of the correlation between two phenomena can be in the form of the following:

Linear Relation	negative Linear Relation
non-Linear Relation	no Relation

In this lesson, you are going to learn how to identify the regression line equation. The purpose of this study is to help the researcher know the type of the given data and perform correct predictions through them.

Remember

- **The function is a relation between the two sets x and y where each element of set x elements has only one element of set y elements**
- **The function is defined whenever each domain, Codomain and rule are known.**

Definition Regression is a statistical method by which the value of a variable can be estimated in terms of the other variable.

It has several types:

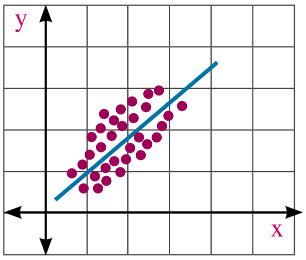
a **simple linear regression:**in which, the dependent variable (y) depends on one variable (x) through a linear relation.

b **multiple regression:**in which, the dependent variable (y) depends on more than an independent variable.

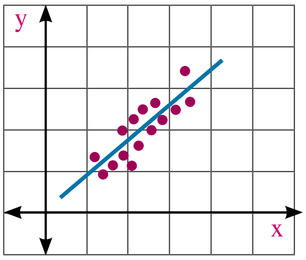
c **Non-linear regression:** if the relation between the dependent variable (y) an nonlinear independent variables (of second degree, third degree, exponential, or logarithmic....).

In this lesson, we concentrate only on the simple linear regression. The following figures illustrate the relation between the value of the correlation relation and the difference of the point positions on the regression line. The closer the points are congruent to this line, the more the value of (r) increases or decreases till all the points get congruent on the line. In this case, the value of (r) is either (+1) or (-1).

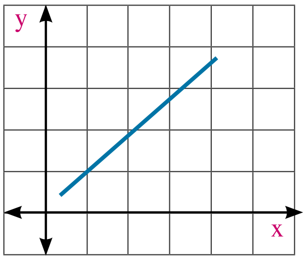
materials: scientific calculator Microsoft Excel



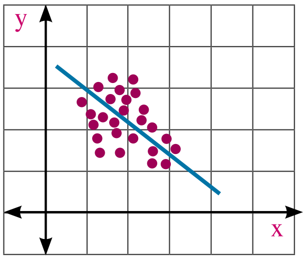
(1) moderate direct correlation



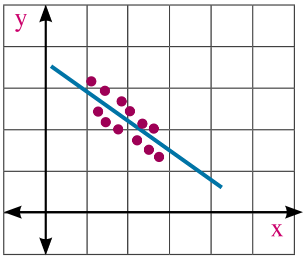
(2) strong direct correlation



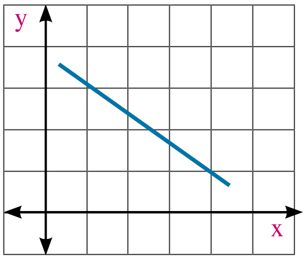
(3) perfect direct correlation



(4) moderate inverse correlation



(5) strong inverse correlation correlation



(6) perfect inverse correlation

Regression line equation:

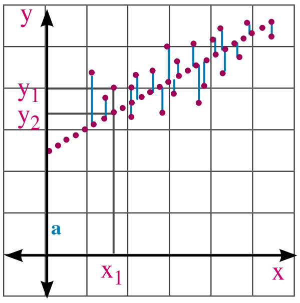
You have previously learned in the analytic geometry the straight line equation whose slope is m and y-intercept of c is $y = mx + c$.

In turn to the scatter diagrams shown previously, we find that if the scatter diagram starts as in figures (2) or (5), this indicates that the relation between the two variables are linear. Since we can imagine in the existence of a straight line, the points are located on and near the straight line even all the points do not lie on it. If the scatter diagram starts as in figures (1) or (4), we doubt about the linearity of the relation between the two variables. As a result, our main mission here is to use the seen pairs of values (x_r , y_r) to find the best straight line matching the points of the set of the sample. Let its equation be:

$$y = a + bx$$

The most common method to find the best values of a and b is called the least square method.

The least square method:



Of the mentioned previously, we knew that in case of correlation, it is not necessary all the points lie on the regression line. Therefore, there is an error ratio for the points which do not lie on the regression line. To get the best regression line, the least possible valued deviations should be reduced (the proper regression line should pass or approach to the greatest number of scattering points). if (x, y) is a true point of data and $(x, \hat{y})$ is the point locating off on the regression line ($\hat{y}$ is read as y hat), then the proper regression line when $|\hat{y} - y|$ is the least for the values of x or when $\Sigma(\hat{y} - y)^2$ is the least and by supposing the regression line equation is $\hat{y} = a + bx$.

The absolute difference = $|(a + bx)-y|$

The required is to determine the values of a,b such that the absolute difference is minimum by solving the following two equations

$\Sigma y = n a + b \Sigma x$ **(1)** $\qquad \Sigma x y = a \Sigma x + b \Sigma x^2$ **(2)**

From equation **(1)** $a = \dfrac{\Sigma y - b \Sigma x}{n}$ And by substituting in **(2)**.

$b = \dfrac{n \Sigma x y -(\Sigma x)(\Sigma y)}{n \Sigma x^2 - (\Sigma x)^2}$ s called the regression coefficient of y on x. It expresses the slope of the regression line on the positive direction of X axis.

The regression line equation of y on x is used for:

1- predicting the value of Y if the value of X is known.

2- identifying the error which can be identified by the relation:

Error = | Table value - the value satisfying the regression equation |

Note: when the regression equation is used for prediction (estimation), it is favorable not to exceed the range of the variable x used in calculating the regression equation much.

Critical thinking: the value of the regression coefficient refers to the correlation. Explain.

Example

(1) The next table illustrates the production of a summer crop (y) of the cultivated land (x) in feddan.

Cultivated land (x) in feddan	50	200	110	80	120	74.5	88.9	5.7	11	3.2
Production of (y) in kg	140	500	400	300	356	240.5	200.6	33.5	69.8	18.7

First: find the regression line equation.

Second: predict the value of the production in kg if the cultivated land is 100 feddans.

Third: find the error in the production if you know that the cultivated land is 120 feddans.

Solution

Solution by using the scientific calculator:

1- enter the data:

Follow up the same method explained previously in Example (1) in the previous lesson (correlation) to enter the data.

2-recall the products(results):

Press the following buttons:

Use the following buttons to find the results of the following operations: SHIFT 1 (STAT) from the drop-down list and press the button Select 3

A new list from 1 to 8 (sums of results) appears and you select as follows:

$$2 : \Sigma x = 743{,}3$$
$$4 : \Sigma y = 2259.1$$
$$1 : \Sigma x^2 = 89017{,}19$$
$$5 : \Sigma xy = 254489{,}18$$

First: we calculate the value of the constant b from the relation:

$$b = \frac{n \Sigma xy - \Sigma x \Sigma y}{n \Sigma x^2 - (\Sigma x)^2}$$

$$= \frac{10 \times 254489{,}18 - 743{,}3 \times 2259{,}1}{10 \times^2 (743{,}3) - 89017{,}19} \simeq 2.5637$$

we calculate the value of the constant a from the relation: $a = \bar{y} - b\bar{x}$

Where : $\bar{x} = \frac{\Sigma x}{n}$, $\bar{y} = \frac{\Sigma y}{n}$

$$\therefore \bar{x} = \frac{743{,}3}{10} = 74.33 \ , \ \bar{y} = \frac{2259{,}1}{10} = 225.91$$

$$\therefore a = 225.91 - 2.5637 \times 74.33 \simeq 35.35$$

Note:

The constant can be directly calculated as follows:

$$\because a = \frac{\Sigma y - b \Sigma x}{n} \qquad \therefore a = \frac{2259{,}1 - (2{,}5637 \times 743{,}3)}{10} \simeq 35.35$$

$\therefore$ The regression line equation is : $\hat{y} = 2.564\,x + 35.35$

Second: From the regression line equation: $\hat{y} = a + b\,x$

$\because \hat{y} = 2.564\,x + 35.35$, $x = 100$

$\therefore \hat{y} = 2.564 \times 100 + 35.35 = 291.72$ Kg

The result can be checked using the calculator as follows:

100 [SHIFT] [1] (STAT) [5] (Reg) [5] : $\hat{y}$ [=]

Third: To find the error in the production if you know that x= 120 feddans:

$\because \hat{y} = 2.564\,x + 35.35$

$\therefore \hat{y} = 2.564 \times 120 + 35.35 \simeq 343$

$\because$ Error = |Table value - the value satisfying the regression equation|

$\therefore$ Error = | 356 – 343| = 13

Activity

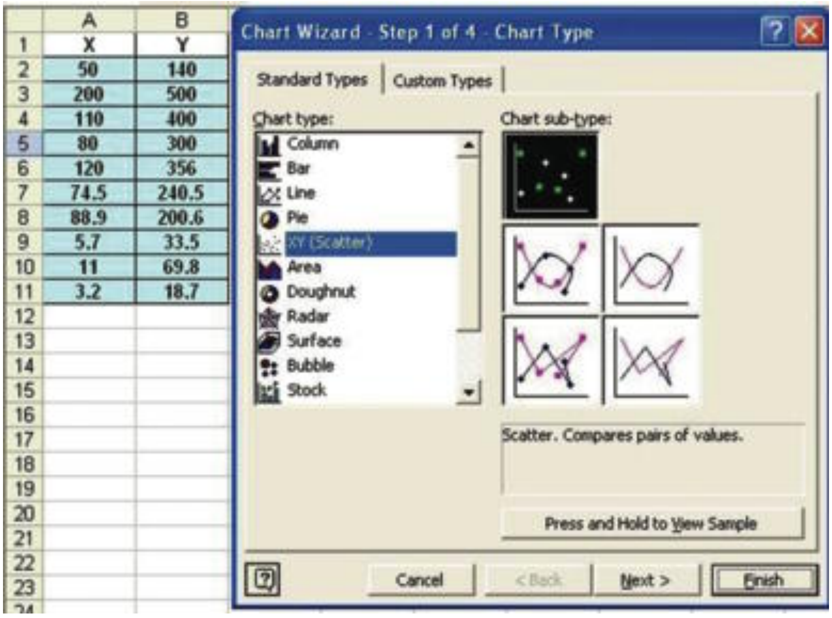
First: check the solution of the example above using (Microsoft Excel).

Second: check the solution of the example above using the statistical program (SPSS).

First: using (Microsoft Excel)

1- start (Microsoft Excel) and enter the previous data in the cells of the two columns a and b under the name x and y as two real variables or the real name of such data as shown in figure 1.

2- from the tool bar, press chartwizard to get chart type, then from the list xy scatter, press finish as in figure 2.



	A	B
1	X	Y
2	50	140
3	200	500
4	110	400
5	80	300
6	120	356
7	74.5	240.5
8	88.9	200.6
9	5.7	33.5
10	11	69.8
11	3.2	18.7

figure (2)

3- figure 3 illustrates the graphical representation of the points listed in the table and is called scatter diagram. We select the figure shaded in black which appears here after changing the background as shown in the figure.

4- the values on the horizontal axis represent the values of X for the data and the vertical axis of the values Y. were, we are going to find the regression line equation which is in the form:

$y = a + bx.$

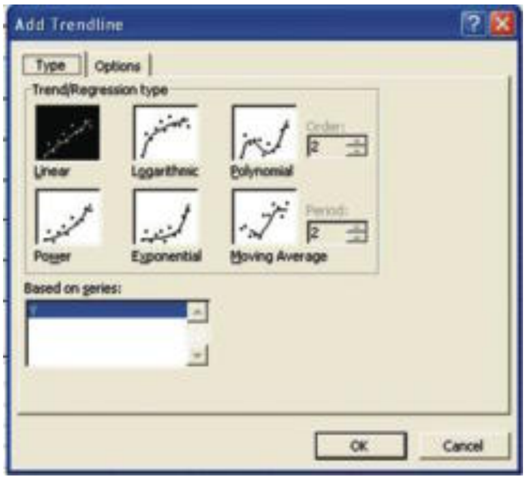


figure 3

5- do a mouse right click on a point (in figure 4). the shown list appears where we select Add Trendline and by doing a mouse click, we get the following figure which appears in six scatter diagrams. We have selected the first scatter diagram as shown by shading in black as a proper choice. Since we need the straight line, we use Option to identify the required by doing a mouse click where the following box appears:

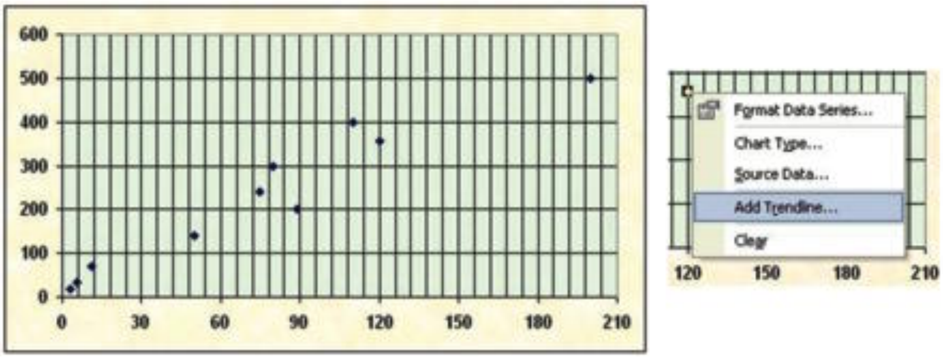


figure (4)

6 - label Display equation on chart as shown in figure 5.

7 - press OK to get the required:

a the figure showing the regression line is the moderate of the points representing the data pairs.

b the regression line equation in figure 6. were, we have moved the equation from its place in the figure above and changed the line to illustrate the matter.

The following figure is the result of the operation and it shows the required especially the following equation:

$Y = 35.35 + 2.5637x$

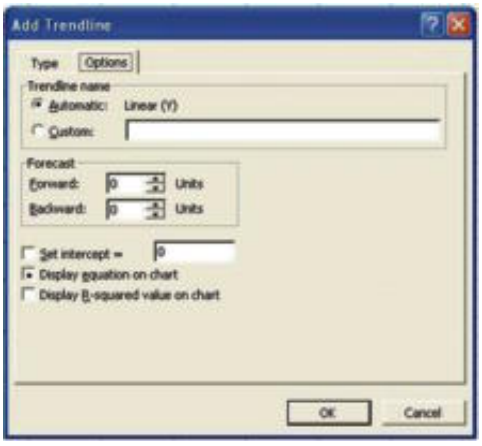


figure (5)

It is a regression line equation and it is the same equation which we have got in the previous solution.

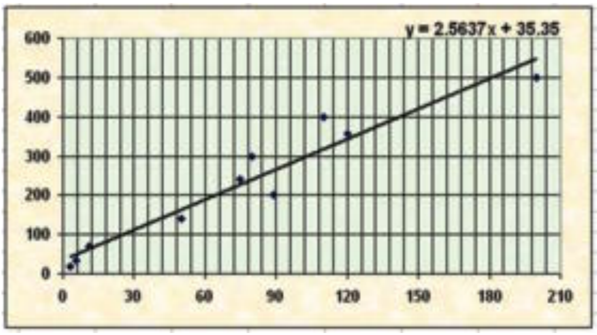
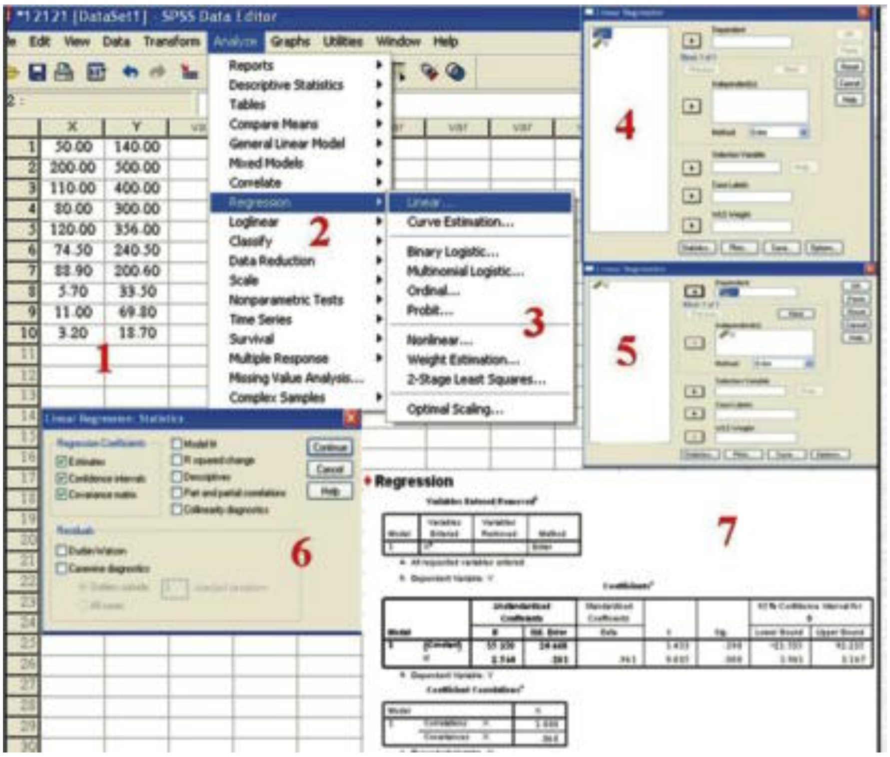


figure (6)

Using SPSS program



(figure 7)

Example

2 Mining: the following table shows the data about the average price of the oil barrel and the rates of the economic growth in a country within eight years. Required is to:

Price of the oil barrel (x)	36	40	36.2	31.1	29.7	16.3	18.7	14.6
Rates of the economic growth(y)	0.91	3.5	3.2	2.7	2.3	– 1	– 0.9	– 1.6

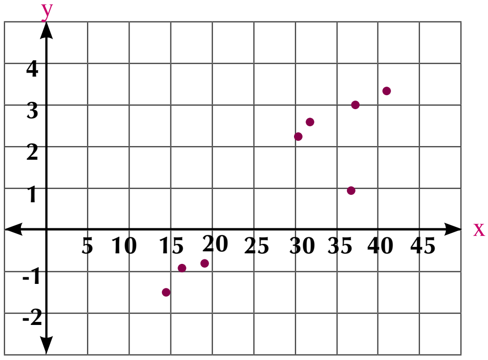
First: graph the scatter diagram and show the type of the correlation.

Second: find the regression line equation of the given data.

Third: predict the economic growth when the price of the oil barrel is \$15, then when the price is \$35.

Solution

First: the figure opposite represents the scatter diagram and it shows also the correlation is direct.



x	y	x^2	y^2	x y
36	0.91	1296	0.8281	32.76
40	3.5	1600	12.25	140
36.2	3.2	1310.44	10.24	115.84
31.1	2.7	967.21	7.29	83.97
29.7	2.3	882.09	5.29	68.31
16.3	– 1	265.69	1	– 16.3
18.7	– 0.9	349.69	0.81	– 16.83
14.6	– 1.6	213.16	2.56	– 23.36
222.6	9.11	6884.28	40.2681	384.39

From the table data:

$\Sigma y = 9{,}11$ $\quad\quad \Sigma x = 222{,}6$

$\Sigma x\, y = 384{,}39$ $\quad\quad \Sigma x^2 = 6884{,}28$

Second: we calculate the value of the constant b from the relation:

$$b = \frac{n\,\Sigma x\,y - \Sigma x\,\Sigma y}{n\,\Sigma x^2 - (\Sigma x)^2}$$

$$= \frac{8 \times 384{,}39 - (222{,}6 \times 9{,}11)}{8 \times^2 (222{,}6) - 6884{,}28} \simeq 0{,}1896$$

$\because a = \dfrac{\Sigma y - b\Sigma x}{n}$ $\quad\quad \therefore a = \dfrac{9{,}11 - (0{,}1896 \times 222{,}6)}{8} \simeq -4.1368$

$\because$ The regression line equation is: $\hat{y} = a + b\,x$

$\hat{y} = 0{,}1896 \;\therefore x - 4{,}1368$

Third:

When $x = 15$ $\quad$ then : $\hat{y} = 0.1896 \times 15 - 4.1368 \simeq -1.2928$

When $x = 35$ $\quad$ then : $\hat{y} = 0.1896 \times 35 - 4.1368 \simeq 2.4992$

Try to solve

1 in a study to find the relation between the income (x) and consumption (y) in thousand pounds, the results show that:

$\Sigma x = 120$, $\Sigma y = 100$, $\Sigma xy = 516$

$\Sigma x^2 = 720$, $\Sigma y^2 = 410$, $n = 40$

a find the linear correlation coefficient between x and y using the Pearson’s method and identify its type.

b find the regression line equation.

c predict the value of consumption (y) when the income is 10000 LE

Exercises 1 - 2

First: choose the correct answer:

1. the statistical equation of the regression line equation where b is the regression coefficient is:

a $\hat{y} = a x + b$ b $\hat{y} = a + b x$

c $\hat{y} = a y + b$ d $\hat{y} = a + b y$.

2. if the regression line equation is: $\hat{y} = 2 + 0.5x$ then the value of y expected when x = 6 is:

a 4 b 5 c 7 d 8

3. if the two points (11.5 , 10) and (6.5 , 5) lie on the regression line y on x, then the correlation between x and y is:

a direct b inverse c perfect d nihilistic

4. if the two points (5 , 13) ,and (14 , 4) lie on the regression line y on x, then all the following points lie on the same line except:

a (15 , 5) b (10 , 8) c (6 , 12) d (5 , 13)

5. if all the points in a scatter diagram lie on a straight line whose slope is negative, then the correlation coefficient between x and y equals:

a 1 b zero c - 0,5 d -1

6. if all the points in a scatter diagram lie on a straight line whose slope is positive, then the correlation coefficient between the two variables equals:

a - 1 b zero c $\frac{1}{2}$ d 1

Second: Answer the following questions:

7. the next table illustrates the relation between the two variables x and y:

x	5	8	10	14	16	20
y	4	6	9	11	12	15

a graph the scatter diagram b find the regression line equation

c predict the value of y when x = 12

8. from the data in the next table:

x	20	33	30	40	13	15	26	25
y	7	8	9	11	4	5	8	9

a predict the value of y when x = 35

b find the error in y if x= 30

9. in a statistical study to find the relation between two variables x and y, we have got the next data:
$n = 10$, $\bar{x} = 8$, $\bar{y} = 10$, $\Sigma xy = 870$, $\Sigma x^2 = 665$, $\Sigma y^2 = 1400$ find:

a) linear correlation coefficient

b) regression line equation

10. if : $\Sigma x = 30$, $\Sigma y = 40$, $\Sigma xy = 162$
$\Sigma x^2 = 210$, $\Sigma y^2 = 304$, $n = 6$ find :

a) regression line equation

b) linear correlation coefficient between x and y and identify its type

11. sales: In a market for selling used cars, the sales were as follows:

car age (x)	3	2	1	1	5	6	1	4
Sale price (y)	54	80	74	98	45	40	85	60

Find:

a) Pearson's linear correlation coefficient

b) Regression line equation

12. Economy: the next table represents the monthly income (x) and the expenditure (y) for a group of families in hundred pounds:

Income (x)	38	27	39	40	56	66	42	44
expenditure (y)	19	25	20	28	31	38	27	22

a) find Pearson's rank correlation coefficient and identify its type.

b) find the regression line equation

c) Estimate the value of the expenditure (y) if the income (x) is 5000 LE

d) find the error in (y) if (x)= 40

13. Family to study the relation between the income (y) and consumption (x) in hundred pounds monthly in a city, a 40- family sample has been taken to give the following results:
$\Sigma x = 100$, $\Sigma y = 120$, $\Sigma xy = 516$, $\Sigma x2 = 410$, $\Sigma y^2 = 720$.

a) find the regression line equation

b) predict the family income whose monthly consumption is 700 LE.

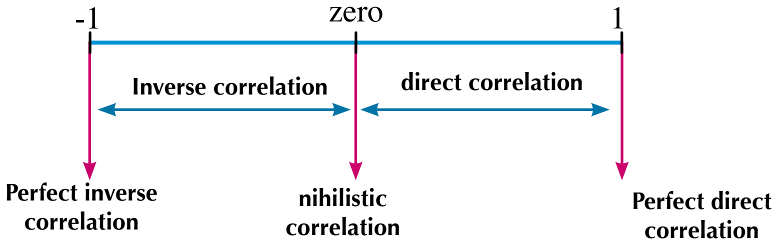
Unit summary

1 Correlation is a statistical method by which the degree and the type between two variables can be determined.

2 Scatter diagram is a graphical representation of a number of the ordered pairs (x,y) to describe the relation between two variables

3 Simple linear correlation is known as the measure of the degree of the relation between two variable

4 Correlation coefficient is denoted by the symbol (r). It is a relatively quantitative scale measuring how much the two variables are correlated where $-1 \leqslant r \leqslant 1$ It is said that the correlation is perfect direct if the correlation coefficient r = 1, is perfect inverse if the correlation coefficient r = -1 and the correlation is not existed when r =0.



5 Pearson's linear correlation coefficient:

$$r = \frac{n\,\Sigma x\,y - \Sigma x \times \Sigma y}{\sqrt{n\,\Sigma x^2 - (\Sigma x)^2}\,\sqrt{n\,\Sigma y^2 - (\Sigma y)^2}}$$

6 Spearman's rank correlation coefficient:

$$r = 1 - \frac{6\Sigma D^2}{n(n^2 - 1)}$$

7 Regression is a statistical method by which the value of a variable can be estimated in terms of the other variable.

8 Regression line equation: $\hat{y} = a + b\,x$

where :

a is the length of y-intercept

b is the regression coefficient of y on x and it expresses the slope of the regression line on the positive direction of X-axis.

$$\mathbf{b} = \frac{n\,\Sigma x\,y - \Sigma x\,\Sigma y}{n\,\Sigma x^2 - (\Sigma x)^2}, \quad a = \frac{\Sigma y - b\Sigma x}{n}$$

9 The regression line equation is used for:

Predicting the value of y if the value of x is known

Identifying the error which can be determined by the relation:

Error = | Table value - the value satisfying the regression equation |

General exercises

First: choose the correct answer:

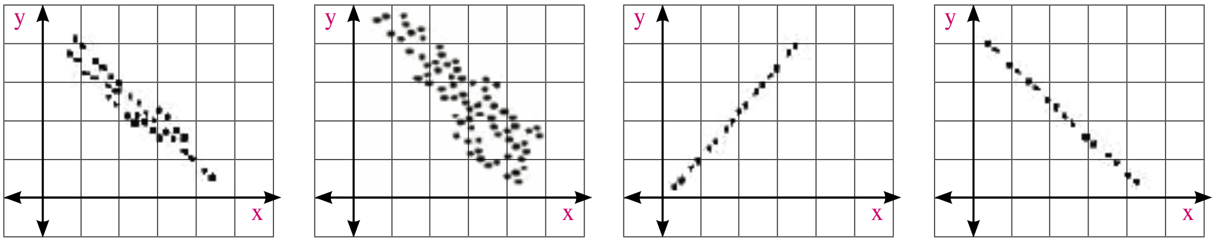
1 The strongest inverse correlation coefficient of the following is:

a – 0.9 b – 0.5 b – 0.1 d zero

2 The smallest correlation coefficient of the following is:

a - 1.1 b - 0.9 b 0.12 d 1.02

3 the figure shows a strong inverse correlation between x and y is figure:



a b c d

4 If the regression line equation is: $\hat{y} = 3 - x$ then the type of the correlation between the two variables x and y is:

a perfect direct b non-correlation c nihilistic d perfect inverse

5 If the regression line equation is: $\hat{y} = 7 - 0.8$, then the value of y expected when x = 5 is:

a 2 b 3 c 5 d 7

6 If the two point (2 , 8) and (7 , 3) lie on the regression line of y on x and the correlation is perfect, then the linear correlation coefficient equals:

a - 1 b zero c $\frac{1}{2}$ d 1

Second: Answer the following questions:

7 If

$\Sigma x = 50$ $\Sigma y = 40$ $n = 10$

$\Sigma xy = 213$ $\Sigma x^2 = 298$ $\Sigma y^2 = 176$

Find the value of Pearson's correlation coefficient between the two variables and identify its type and degree.

8 From the data of the following table:

x	4	5	6	7	8	9	10
y	12	7	12	3	14	15	21

calculate Pearson's correlation coefficient between the two variables and identify its type.

9 From the data of the following table:

x	32	42	40	35	31	46	50	33
y	25	34	35	30	17	28	42	19

Calculate Spearman's rank correlation coefficient between the values of x and y and identify its type.

10 From the data of the following table:

x	excellent	good	Very good	Pass	weak	good
y	good	weak	pass	excellent	Very good	pass

Calculate Spearman's rank correlation coefficient between x and y.

11 Trade: the next table illustrates the sales x and the profits y for a 6- company group. Required is to calculate Pearson's correlation coefficient between the size of sales and the profits.

Sales x	500	600	400	480	550	100
Profits y	300	400	250	200	400	90

12 From the data of the following table:

x	10	12	15	12	14	8
y	6	8	6	6	9	5

a Calculate Pearson's linear correlation coefficient between the two variables and identify its type.

b Find the regression line equation, then predict the value of $\hat{y}$ When x =7.

13 Trade:

To study the relation between the quantity (y) of an item and the price (x) in LE, we have got the next data:

$\Sigma x = 49$, $\Sigma y = 77$, $\Sigma xy = 609$, $\Sigma x^2 = 371$, $\Sigma y^2 = 1049$, n = 7 find :

a Pearson's linear correlation coefficient between the quantity required and the price:

b Estimate the quantity ($\hat{y}$) When the price is 21 LE.

14 Sports: the next tables illustrates the age of a person and the number of training hours.

Age	20	25	33	37	50	58
number of training hours	10	6	3	2	1.5	1

a Find the regression line equation

b Predict the number of training hours when this person is 40 years old.

c Calculate the error when this person is 33 years old.

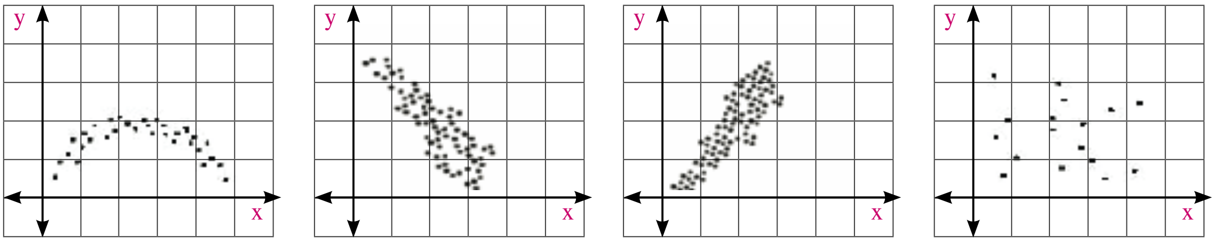
For more activities and exercises www.sec3mathematics.com.eg

Accumulative test

Short-answerd questions

First: choose the correct answer (From 1 to 10):

1. the sum of the values whose mean is 8 and their number 7 equal:
   a 40 b 56 c 60 d 80
2. the figure representing a direct relation between x and y is:



   a b c d

3. The relation between the square side length and its area is a ....correlation:
   a strong direct b strong inverse c perfect direct d perfect inverse
4. If the two variables increase or decrease together, the correlation between them is
   a direct. b inverse. c non-linear. d nihilistic.
5. The correlation coefficient of a digital scale whose value ranges between
   a [0 , 1] b]- 1 , 1[c [- 1 , 1] d [- 1 , 1] – {0}
6. The variable required to be estimated in the regression line equation is called a (n) .... variable.
   a independent b dependent c direct d inverse

Second: Give the reason:

7. The sign of regression coefficient refers to the type of correlation (direct or inverse). Explain.

Long- answered questions:

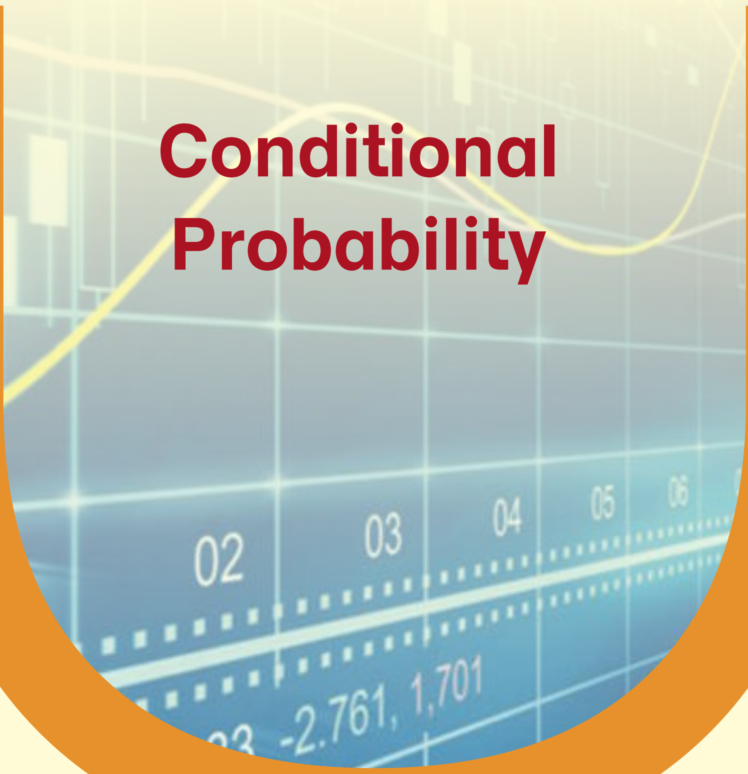
8. From the data in the opposite table

x	65	66	67	67	68	69
y	70	72	68	64	68	67

   a Find the linear correlation coefficient between the two variables x and y and identify its type.
   b Predict the value of $\hat{y}$ when x = 62
   c Calculate the error in y if x = 66
9. The next table illustrates the relation between a driver's age and the number of the traffic tickets he got in a year.

Driver's age (x)	45	28	52	24	56	32	63	38
number of traffic tickets (y)	4	7	2	7	2	5	1	3

   a Find the rank correlation coefficient between x and y
   b Find the regression line equation of y on x c Estimaten $\hat{y}$ if x = 40
   d Calculate the error in y if x = 38



Unit 2

Conditional Probability

introduction

We have previously known that statistics is a branch of mathematics concerning with collecting, ordering and interpreting data in order to get the ability to take the proper decisions for a certain phenomenon. The probabilities are considered the mathematical background for the statistical methods. Long ago, the researchers used the probabilities for social, economic and healthy reasons.

The probability has been founded as it is today by a great number of scientists such as the French scientist (Pierre Simon Laplace- 1749- 1827), the English scientists (De Morgan 1806- 1871), (John Vin 1834-1923) and the Russian scientist (Andrei Markov 1856-1922) and others.

Pierre Simon Laplace

De Morgan

John Vin

Andrei Markov

It is worth mentioning that the applications of statistics and probability are much in the educational, social and economic fields. In this unit, we are going to learn the conditional probability between two events, its theory and its applications in the different daily life situations. Besides, we will learn the independent and dependent events.

Unit objectives

By the end of the unit and carrying out the involved activities, the students should be able to.

- Identify the mutually exclusive events and events are not mutually exclusive
- Identify the conditional probability
- Deduce theories on the conditional probability
- Identify the independent and dependent events
- Apply the conditional probability in the different daily life situations

Key Terms

- *Mutually Exculusive events*
- *Events are not mutually exclusive*
- *Conditional probability*
- *Independent Events*
- *Dependent Events*

Lessons of the unit

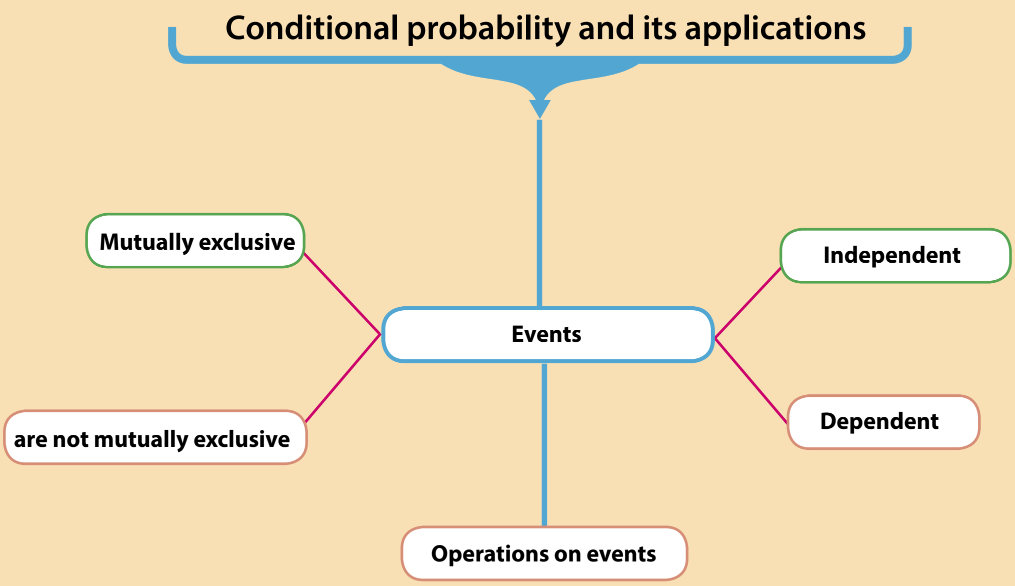
Lesson (2-1): Conditional probability

Lesson (2-1): Independent events.

Materials

Scientific Calculator

Unit planning organizer



Unit Two

2 - 1 Conditional Probability

You will learn

- Mutually exclusive events
- Events are not mutually exclusive
- Conditional probability

Key terms

- Mutually Exclusive Events
- Events are not Mutually Exclusive
- Conditional probability

Introduction:

You have previously learned to calculate the probability of an event (let it be A) of a random experiment by knowing the relation between the number of the elements of this event n(A) and the number of the elements of the sample space of the random experiment n(S) through the relation

$$P(A) \text{ (the probability of occurring even A)} = \frac{\text{the number of outcomes in event } n(A)}{\text{the number of outcomes in sample space } n(s)}$$

Mutually Exclusive Events

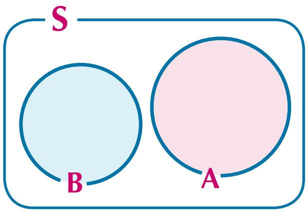
Of your study to the probability, you have learned that the mutually exclusive events are the events which cannot occur at the same time because when an event occurs, the other events are hindered to occur. This means that there are not common elements of the elements forming them.

Mutually exclusive events:

They are the two events which do not share in any element and their intersection is the null set ϕ.

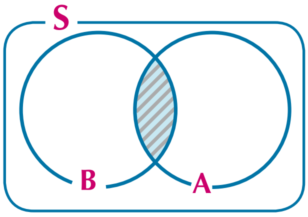
If A and B are two mutually exclusive events, then $A \cap B = \phi$

$\therefore P(A \cap B) = \text{Zero}$ and $P(A \cup B) = P(A) + P(B)$



Events are not Mutually Exclusive

They are the two events in which if an event occurs, it does not trap the occurrence of the other(there are common elements between them).Then,



(1) $P(A \cup B) = B(A) + P(B) - P(A \cap B)$

(2) $P(A') = 1 - P(A)$

(3) $P(A - B) = P(A) - P(A \cap B)$

(4) $P(A \cap B') = P(A - B) = P(A) - P(A \cap B)$

(5) $P(A' \cap B) = P(B - A) = P(B) - P(A \cap B)$

Materials

- Scientific calculator - Computer graph program.

Conditional Probability

If A and B are two events from S, sometimes the information is available that an events such as B has been occurred P (B). in this case, the occurrence of the event B may affect the probability of occurring the event A. The probability of occurring the event A in condition of occurring the event B can be calculated by knowing the relation among the outcomes of the event A or the outcomes of the event B.

Introductory example: In an experiment of rolling a regular dice once, the sample space S is : S = {1 , 2, 3, 4, 5, 6} , If the event A = {1 , 2, 3} is the event of the appearance of a number less then 4

It becomes clear that: $P(A) = \frac{n(A)}{n(S)} = \frac{3}{6} = \frac{1}{2}$

If the event B = {2, 4, 6} is the event of the appearance of an even number.

We ask: What is the probability of occurring the event A, if we know that the event B has already occurred?

In other words: what is the probability of getting an even number less than 4?

We notice that the given condition reduces the sample space into set B = {2, 4, 6}

Then, the event matching the appearance of an even number is $A \cap B = \{2\}$

And the probability required is: $\frac{P(A \cap B)}{P(B)} = \frac{1}{6} \div \frac{1}{2} = \frac{1}{3}$

This example shows us how the probabilities of some events differ in regard to the difference of the sample space.

Learn

Conditional Probability

If S is the sample space of a random experiment and A and B are two events of this sample space, then the probability of occurring the event A in condition of occurring the event B is denoted by the symbol P(A|B) and read as the probability of occurring the event A in condition of occurring the event B. it can be determined by the next relation:

$$P(A | B) = \frac{P(A \cap B)}{P(B)} \quad \text{where } P(B) > 0$$

Notice that: the conditional probability has the same properties of the unconditional one.

1- $0 \leqslant P(A | B) \leqslant 1$

2- $P(S | B) = \frac{P(S \cap B)}{P(B)} = \frac{P(B)}{P(B)} = 1$

3- If $A_1 \cap A_2 = \phi$ then, $P[(A_1 \cup A_2) | B] = P(A_1 | B) + P(A_2 | B)$

Regard the following:

- $P(A | B) \neq P(B | A)$
- $P(A' | B) = 1 - P(A | B)$
- $P(A \cap B) = P(A|B) \times P(B)$ in a condition $P(B) > 0$
- $P(A \cap B) = P((B | A) \times P(A)$ in a condition $P(A) > 0$

Example the conditional probability

1 A regular die has been rolled once. calculate the probability of appearing the number 2 known that the number appeared is even.

Solution

Let the sample space S= { 1, 2, 3, 4, 5, 6} , A = {2} , B = { 2 , 4, 6}

Then, $P(B) = \frac{3}{6} = \frac{1}{2}$, $P(A \cap B) = P(A) = \frac{1}{6}$

$\because P(A | B) = \frac{P(A \cap B)}{P(B)}$

$\therefore P(A | B) = \frac{1}{6} \div \frac{1}{2} = \frac{1}{6} \times 2 = \frac{1}{3}$

The probability of appearing the number 2 known that the number appeared is even is $\frac{1}{3}$

> **Notice**
>
> **In the conditional probability, notice that the event after the phrase (What is the probability of) is the events which we start with and the event next to the phrases (known that A, if A, if known that a and so on) is the conditional event.**

Try to solve:

1 A regular die has been rolled twice. What is the probability the number of points in the first roll is not more than 4 if you know that the absolute difference between the two number appeared equals 2?

Example Doing the operations

2 If A and B are two events of the sample space where P(A) = 0.45 , P (B) = 0.6 , P (B | A) = 0.8 find:

a $P(A \cap B)$ b $P(A \cup B)$ c $P(A | B)$ d $P(B' | A)$

Solution

a $\because P(B | A) = \frac{P(B \cap A)}{P(A)}$

$\therefore 0.8 = \frac{P(A \cap B)}{0.45}$ $\therefore P(A \cap B) = 0.8 \times 0.45 = 0.36$

b $\because P(A \cup B) = P(A) + P(B) - P(A \cap B)$

$\therefore P(A \cup B) = 0.45 + 0.6 - 0.36 = 0.69$

c $P(A | B) = \frac{P(A \cap B)}{P(B)} = \frac{0.36}{0.6} = 0.6$

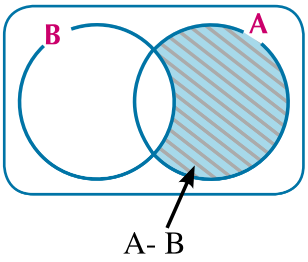
Notice that: $P(A | B) \neq P(B | A)$

d $P(B' | A) = \frac{P(B' \cap A)}{P(A)} = \frac{P(A - B)}{P(A)}$

$= \frac{P(A) - P(A \cap B)}{P(A)}$

$= \frac{0.45 - 0.36}{0.45} = 0.2$

> **Remember**
>
> $P(B \cap A) = P(A \cap B)$
>
> $P(A \cup B) = P(A) + P(B) - P(A \cap B)$
>
> $P(A - B) = P(A) - P(A \cap B)$



Try to solve

2 If A and B are two events of the sample space of a random experiment S where P(A) = 0.7 , P(B) = 0.25 , P(A - B) = 0.45 find:

a $P(A \mid B)$

b $P(B \mid A)$

c $P(A' \mid B)$

d $P(A' \mid B')$

Example Harmonic tables

3 From the data in the next table:

Case	Number of people	
	Wear glasses	Do not wear glasses
Man	800	600
Woman	400	200

Find the probability of a woman wearing glasses has been randomly chosen

Solution

Let n(S)= Number of people under study=2000,

A is the event that the person chosen is a woman,

B is the event that the person chosen wears glasses.

$$P(A \cap B) = \frac{400}{2000} = \frac{1}{5}$$

$$P(B) = \frac{1200}{2000} = \frac{3}{5}$$

Required is to find the probability of A known that B has already occurred. i.e. $P(A \mid B)$

$$\because \; P(A \mid B) = \frac{P(A \cap B)}{P(B)}$$

$$\therefore \; P(A \mid B) = \frac{1}{5} \div \frac{3}{5} = \frac{1}{3}$$

The probability of a woman wearing glasses has been randomly chosen is $\frac{1}{3}$

Try to solve

3 In the previous example, find the probability that:

a A man does not wear glasses is randomly chosen.

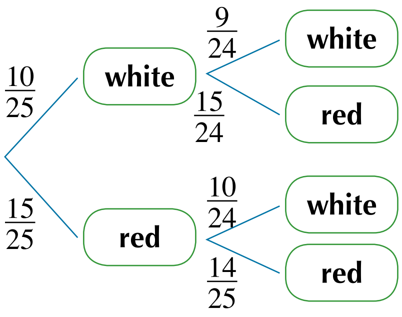
b A woman or man wearing glasses is randomly chosen.

Example — Tree diagram

4 A bag contains 10 white balls and 15 red balls. Two balls have been consecutively drawn without replacing. What is the probability the two drawn balls are white?

Solution

In this example, we notice that drawing the balls has been conducted consecutively. As a result, it is subjected to the order. In other words, the second drawing is conditioned by the occurrence of the first drawing. This example can be represented by the tree diagram as shown in the figure opposite.



Let A denote the event that the first drawn ball is white and B denote the event that the second drawn ball is white.

Let (B | A) denotes the event of drawing the second ball in a condition the first ball has already been drawn .

and **let** (A ∩ B) denotes the event of drawing two white balls..

$\because$ $P(B \mid A) = \dfrac{P(B \cap A)}{P(A)}$

$\therefore$ $\dfrac{9}{24} = \dfrac{(A \cap B)}{\frac{10}{25}}$

$\therefore P(A \cap B) = \dfrac{9}{24} \times \dfrac{10}{25} = \dfrac{3}{20}$

The probability the two drawn balls are white is $\dfrac{3}{20}$

Try to solve

4 In the previous example, find the probability the two drawn balls are red.

Example — Education

5 100 learners are studying in a language institute. The number of learners studying English is 60, the number of the learners studying French is 50 and the number of the learners studying both languages is 35. If a learner has been randomly chosen from the institute, find the probability the learner studies:

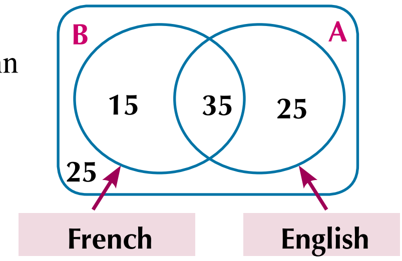
a a language at least.

b English if he (she) studies French.

c French if he (she) studies English.

Solution:

The data of this problem can be explained using the Venn diagram as shown in the figure opposite.



Let the following events be as follows:

The learner studies English = A

The learner studies French = B, then:

$P(A) = \frac{60}{100} = 0.6$, $P(B) = \frac{50}{100} = 0.5$, $P(A \cap B) = \frac{35}{100} = 0.35$

a The probability the learner studies a language at least is $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

$\therefore P(A \cup B) = 0.6 + 0.5 - 0.35 = 0.75$

I.e. the probability the learner studies a language at least is 0.75

b the probability the learner studies English if he (she) studies French = $P(A \mid B)$

$\because P(A \mid B) = \frac{P(A \cap B)}{P(B)}$

$\therefore P(A \mid B) = \frac{0.35}{0.5} = 0.7$

I.e. the probability the learner studies English if he (she) studies French is 0.7

c the probability the learner studies French if he (she) studies English = $P(B \mid A)$

$\because P(B \mid A) = \frac{P(B \cap A)}{P(A)}$

$\therefore P(B \mid A) = \frac{0.35}{0.6} \simeq 0.583$

I.e. the probability the learner studies French if he (she) studies English is 0.583 approximately.

Try to solve

5 Two players A and B shoot at the same time towards the goal. If the probability the player A scores a goal = $\frac{2}{5}$, the probability the player B scores a goal = $\frac{1}{4}$. and the probability both players A and B score together = $\frac{1}{6}$, Find the probability of:

a Scoring a goal.

b Scoring a goal by player A if player B scored a goal.

c Scoring a goal by player B if player A scored a goal.

Exercises 2 - 1

First: Choose the correct answer:

1. in an experiment for tossing a regular coin twice, the probability of appearing Tail in the second toss if Head appears in the first toss equals:

   a) $\frac{1}{4}$ b) $\frac{1}{2}$ c) $\frac{3}{4}$ d) 1

2. in an experiment for rolling a regular die once, the probability of appearing a prime even number greater than 1 is:

   a) $\frac{1}{5}$ b) $\frac{2}{5}$ c) $\frac{3}{5}$ d) $\frac{4}{5}$

3. in an experiment for rolling a regular die once, the probability of appearing the number 3 known that the appearing number is odd is:

   a) $\frac{1}{4}$ b) $\frac{1}{3}$ c) $\frac{1}{2}$ d) $\frac{3}{4}$

4. If $P(A \cap B) = \frac{2}{5}$, $P(A) = \frac{4}{5}$, then $P(B \mid A) =$

   a) $\frac{1}{2}$ b) $\frac{8}{25}$ c) $\frac{1}{4}$ d) $\frac{2}{5}$

5. If $P(A \mid B) = \frac{1}{3}$, $P(B) = \frac{12}{25}$ then $P(A \cap B) =$

   a) $\frac{4}{25}$ b) $\frac{1}{4}$ c) $\frac{25}{36}$ d) $\frac{16}{25}$

Second: Answer the following questions:

6. If A and B are two events of the sample space of a random experiment S where $P(A) = 0.4$, $P(B) = 0.7$, $P(A \mid B) = 0.3$ find:

   a) $P(A \cap B)$ b) $P(A \cup B)$ c) $P(B \mid A)$ d) $P(A \mid B')$

7. If $P(A') = 0.4$, $P(B) = 0.5$, $P(A \cup B) = 0.8$ find $P(A \mid B')$

8. If $P(B \mid A) = \frac{2}{3}$, $P(B \mid A') = \frac{4}{7}$, $P(A) = \frac{3}{5}$ find

   a) $P(A \cap B)$ b) $P(A \cup B)$

9. A die has been rolled once. Calculate the probability the appearing number is a prime number in a condition the appearing number is an odd number.

10. in an experiment of rolling two different dices once, find the probability:

    a) the appearing number on the second die equals 4 known that the appearing number on the first die equals 2.

    b) the sum of the two appearing numbers is even known that the appearing number on the first die equals 6.

11. If the probability of a student to succeed in an exam is 0.7 and the probability to travel abroad if he succeeded is 0.6. What is the probability of his success and traveling abroad?

12 A 45- student class; 27 students study French, 15 students study German and 9 students study both languages. A student is randomly chosen, calculate the probability the chosen student studies:

a a language at least.

b French if he (she) studies German.

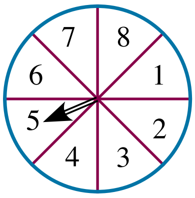
c German if he (she) studies French.

13 two different dices have been rolled once, find the probability of the following events:

a the appearance of number 2 on the two faces together known that the same number appeared on both of them.

b the appearance of number 5 on the two faces together known that each of the appearing number is greater than 4.

c the non appearance of number 3 on any of the two faces known that the two numbers appeared are odd..



14 Spinning wheel: A spinning wheel has been divided into 8 equal circular sectors from 1 to 8. What is the probability the gage lands on the number 5 if known that it has previously landed on an odd number?

15 the next table illustrates the numbers of the sports teams participating in different games:

Sports game	Hand ball	Soccer	Volleyball	Basketball	Hockey ball
Numbers of participating teams	4	10	6	7	3

A game is randomly chosen. What is the probability the game is:

a Hockey ball known that it is not of Volleyball games?

b Basketball known that it is neither of soccer nor Hand ball?

16 a random sample made up of 30 male students and 20 female students to answer some questions related to economy and consuming the energy has been chosen and their answers have been as follows:

Answer	Yes	No	Uncertain	Total
Male students	20	6	4	30
Female students	15	3	2	20

A student is randomly chosen. What is the probability the student chosen is a female and her answer is “Yes“?

17 A box contains 5 white balls and 7 black balls. If 2 balls have been consecutively drawn without replacing, find the probability::

a the second ball is white if the first one is white.

b the first ball is white and the second one is white.

c the second ball is black and the first one is white.

18 Karim and Ziad compete for the presidency of the student Union at their school in three classes, and the following table illustrates the votes obtained by each of them:

	First class	Second class	Third class	Total
Karim	196	174	130	500
Ziad	240	165	135	540

A student is randomly chosen. What is the probability the student chosen::

a voted for Karim known that the student is from the third class.

b voted for Ziad known that the student is from the second class.

19 A job has been announced for and 100 people have applied for this job. If their data have been ordered as follows::

Qualified			Unqualified		
	Married	**Single**		**Married**	**Single**
Male	40	10	**Male**	3	12
Female	10	10	**Female**	10	5

a calculate the probability the employee chosen is married in a condition he (she) is qualified.

b calculate the probability the employee chosen is married and qualified.

c calculate the probability the employee chosen is married in a condition he(she)is unqualified.

20 In the final year exam, 30% of the students have failed in chemistry, 20% have failed in physics and 15% have failed in both chemistry and physics. A student has been randomly chosen.

a what is the probability the student fails in physics if he (she) has already failed in chemistry?

b what is the probability the student fails in chemistry if he (she) has already failed in physics?

c find the probability the student fails in chemistry in a condition he (she) fails in physics.

d find the probability the student succeeds in physics in a condition he (she) succeeds in chemistry.

21 Activity: Use Venn diagram:

A and B are two events of the sample space S where $P(A) = 0.7$, $P(B) = 0.4$, $P(A \cap B) = 0.2$

a represent the previous sets using Venn diagram and write down the probability of these events to occur on the graph.

b find the probability of the following events:

First: the occurrence of event A in a condition event B do not occur.

Second: the occurrence of event B in a condition event A do not occur..

Independent Events

Unit Two

2 - 2

You will learn

- Independent events
- Dependent events

Key terms

- Independent Events
- Dependent Events

Think and discuss

Meditate the following examples

1- tossing a coin and rolling a die once.

2- a student has succeeded in mathematics and chemistry.

3- a ball has been randomly drawn from a box containing 10 balls, then it has been turned back to the box and another ball has been drawn.

4- a student has succeeded in the physics lab-exam and in physics.

5- a ball has been randomly drawn from a box containing 10 balls without turning it back, then another ball has been drawn.

What do you notice?

From the first three examples, we notice that:

1- the outcomes of the coin do not affect the outcomes of the dice .

2- the success or failure of the student in mathematics does not affect his (her) success or failure of the student in chemistry

3 - returning the first ball to the box after drawing it does not change the number of the balls, so the first draw does not affect the second one.

As a result, the events in each example of the first previous three examples are called the independent events.

4- the success of the student in the physics lab-exam affects the success in physics.

5- when a ball has been drawn from the box without returning to the box it affects the number of the balls in the box and so the first draw affects the second draw.

As a result, the events in examples (4) and (5) are called the dependent events.

Learn independent events

Definition **It is said that A and B are two independent events if and only if $P(A \cap B) = P(A) \times P(B)$.**

I.e. the probability of occurring two independent events together equals the probability of occurring the first event multiplied by the probability of occurring the second event

Materials: Scientific calculator- computer graphs.

It is noticed that if the two events A and B are independent and $P(B) \neq 0$

then $P(A \mid B) = P(A)$ **I.e.** the occurrence of an event does not affect the occurrence of the other event.

For example: a regular coin has been tossed twice and the frequency of occurring the Tail and Head has been noticed , then: S = { (H, H), (H, T), (T, H) , (T, T)}

So the probability of any of the outcomes $= \frac{1}{4}$

Let event A represent the appearance of tail in the second time = {(H , T) , (T , T) }

and event B represent the appearance of Head in the first time = {(H , H) , (H , T)}

then $P(A \mid B) = \frac{P(A \cap B)}{P(B)} = \frac{\frac{1}{4}}{\frac{1}{2}} = \frac{1}{2} = P(A)$

I.e. the occurrence of event B does not affect the probability of occurring event A. In other words, the probability of event A does not depend on knowing that the event B is occurred or not. Thus, we say that the two events A and B are independent.

Notice that: The mutually exclusive events A and B are independent if and only if $P(A) \times P(B) = 0$

In other words, if and only if the probability of A or the probability of B equals zero.

Example

1. what is the probability of appearing Head and number 5 in an experiment of tossing a coin once, then rolling a dice?

Solution

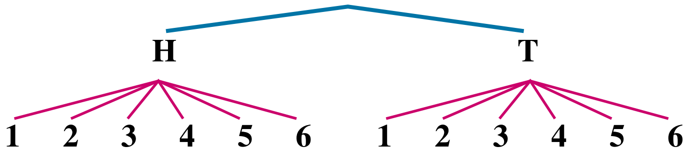
The tree diagram can be used to write down the sample space. We notice that when the coin is tossed, it does not affect the sample outcomes of rolling the dice. This means that the two events are independent.

Let A event of appearing Head, then $P(A) = \frac{1}{2}$, B = event of appearing number 5, then $P(B) = \frac{1}{6}$

$\because P(A \cap B) = P(A) \times P(B)$ $\therefore P(A \cap B) = \frac{1}{2} \times \frac{1}{6} = \frac{1}{12}$

$\therefore$ The probability of appearing Head and number 5 is $\frac{1}{12}$

Note: The probability of appearing Head and number 5 can be found directly by writing down the sample space as shown in the figure opposite.



S = { (H , 1) , (H , 2) , (H ,3) , (H , 4) , (H , 5) , (H , 6) , (T , 1) , (T, 2) , (T, 3) , (T , 4), (T , 5) , (T , 6) }

The event of appearing Head and number 5 = {(H , 5)}

and The probability of appearing Head and number 5 = $\frac{1}{12}$

Try to solve:

1. in the previous example, find the probability of appearing Tail and a prime number.

Example

2 If A and B are two events of a sample space of a random experiment S and P(A)=0.5, P(B)=0.6 and $P(A \cup B) = 0.8$. Explain if A and B are two independent events.

Solution

$\because P(A \cap B) = P(A) + P(B) - P(A \cup B)$

$\therefore P(A \cap B) = 0.5 + 0.6 - 0.8 = 0.3$ **(1)**

$\because P(A) \times P(B) = 0.5 \times 0.6 = 0.3$ **(2)**

From **(1)** and **(2)**, A and B are two independent events.

Notice that: To show the difference between two mutually exclusive events and two independent events, we study the next example:

We know that when we toss a regular coin once, the sample space S={H, T}

We also know that $P(H) = \frac{1}{2}$ and $P(T) = \frac{1}{2}$

Furthermore, we know that the two events H and T are two mutually exclusive events because the occurrence of an event negates the occurrence of other event.

$\therefore P(H \cap T) = \text{zero}$, $\therefore P(H \cap T) \neq P(H) \times P(T)$

I.e. H and T are two mutually exclusive events but they are dependent

Try to solve

2 If A and B are two events of a sample space of a random experiment S where S= {1,2,3,4,5,6} A = {2, 3, 5, 6} and B {1, 4, 5, 6}. Are A and B two dependent events? Explain.

Example

3 Insurance: a man and his wife have insured their life at a life insurance company. If the company has estimated the probability that the man will live more than 20 years to be 0.2 and the probability that his wife will live more than 20 years to be 0.3, find the probability that:

a The man and his wife will live more than 20 years together.

b At least one of them will live more than 20 years.

c Only one of them will live more than 20 years.

Solution

Let A be the event that the man will live more than 20 years $\therefore P(A) = 0.2$,

B the event that the wife will live more than 20 years $\therefore P(B) = 0.3$

a The probability that the man and his wife will live more than 20 years together = $P(A \cap B)$

$\because P(A \cap B) = P(A) \times P(B)$ $\therefore P(A \cap B) = 0.2 \times 0.3 = 0.06$

b The probability that at least one of them will live more than 20 years = $P(A \cup B)$

$\because P(A \cup B) = P(A) + P(B) - P(A \cap B) \therefore P(A \cup B) = 0.2 + 0.3 - 0.06 = 0.44$

c The probability that only one of them will live more than 20 years = P (A ∪ B) - P (A ∩ B)

$\therefore P(A \cup B) - P(A \cap B) = 0.44 - 0.06 = 0.38$

Try to solve

3 Shooting: two soldiers A and B have shot a missile in the direction of a target. If the probability soldier A shot the target is 0.6 and the probability soldier B shot the same target is 0.5, find the probabilities of the following events:

a Shooting the target by soldier A and soldier B together.

b Shooting the target by a missile at least.

c Shooting the target by one missile only. **d** No one Shot the target.

Example

4 Drawing with replacing: A bag contains 6 blue balls and 4 red balls. A ball is randomly drawn, then it is turned back to the bag, then another ball is drawn. What is the probability

a the two balls are red in the two times **b** the two balls are blue in the two times?

c the first ball is red and the second is blue? **d** one ball is red and the other is blue?

Solution

a As long as the process of drawing the ball is accompanied with replacing, the two events are **independent.**

Let S= Sample space, A= drawing the ball first time and B= drawing the ball second time.

$\therefore n(S) = 10, \ P(A) = \frac{4}{10}, \ P(B) = \frac{4}{10}$ **(since the drawing is with replacing)**

$\because P(A \cap B) = P(A) \times P(B)$ $\therefore P(A \cap B) = \frac{4}{10} \times \frac{4}{10} = \frac{16}{100} = \frac{4}{25}$

Similarly:

b The probability the two balls are blue in the two times $= \frac{6}{10} \times \frac{6}{10} = \frac{36}{100} = \frac{9}{25}$

c The probability the first ball is red and the second is blue $= \frac{4}{10} \times \frac{6}{10} = \frac{24}{100} = \frac{6}{25}$

d The probability one ball is red and the other is blue = the probability the first ball is red and the second is blue + the probability the first ball is blue and the second is red

$$= \frac{4}{10} \times \frac{6}{10} + \frac{6}{10} \times \frac{4}{10} = \frac{12}{25}$$

Try to solve

4 The probability of rising the stock market index in a country (A) equals 0.84 and the probability of rising the stock market index in a country (B) equals 0.75. What is the probability the index of the two stock markets of the two countries A and B rises?

Learn — Dependent events

A and B are two dependent events if: $\mathbf{P(A \cap B) \neq P(A) \times P(B)}$

because we know from the conditional probability that:

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)} \qquad P(B) \neq 0$$

$$P(B \mid A) = \frac{P(A \cap B)}{P(A)} \qquad P(A) \neq 0$$

I.e. it can be written as $P(A \cap B) = P(A \mid B) \times P(B)$

$= P(B \mid A) \times P(A) \qquad P(A) \neq 0, P(B) \neq 0$

In other words, the two events A and B are dependent if the probability of the occurrence of one of them affects in a way the probability of the occurrence of the other event.

Probability of dependent events:

Example

5 If S is the sample space of a random experiment where S = {1, 2, 3, 4, 5, 6, 7, 8}, A = {1, 2, 4, 8} and B ={2, 5, 6, 7}. Are A and B two independent events? Explain.

Solution

$\because n(A) = 4 \qquad \therefore P(A) = \frac{4}{8} = \frac{1}{2} \quad , \qquad \because n(B) = 4 \qquad \therefore P(B) = \frac{4}{8} = \frac{1}{2}$

$\because A \cap B = \{2\} \qquad \therefore P(A \cap B) = \frac{1}{8}$ **(1)**

$\because P(A) \times P(B) = \frac{1}{2} \times \frac{1}{2} = \frac{1}{4}$ **(2)**

From **(1)** and **(2)**, $P(A \cap B) \neq P(A) \times P(B)$ so, A and B are two dependent events.

Try to solve:

5 In the previous example If C = {2, 3, 4, 7}. Are B and C two independent events? Explain.

Drawing without replacing:

Example

6 A bag contains 6 blue balls and 4 red balls. If two balls are drawn one after another without replacing. What is the probability:

a the two balls are red?

b the two balls are blue?

c the first ball is red and the second is blue?

Solution

This example is similar to example (3) but the difference is that drawing the balls is conducted without replacing. So, the two events are dependent.

a if the two drawn balls are red, then

The probability the first drawn ball is red and the second is red =

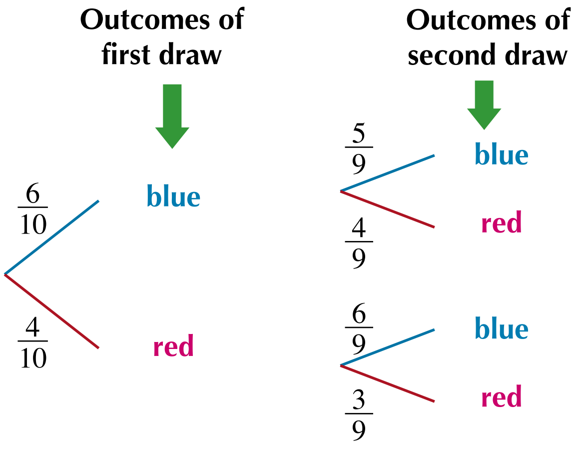
the probability the second drawn ball is red after drawing the first red ball × The probability the first drawn ball is red

$= \frac{4}{10} \times \frac{3}{9} = \frac{2}{15}$

b) if the two drawn balls are blue, then: The probability the first drawn ball is blue and the second is blue = $\frac{6}{10} \times \frac{5}{9} = \frac{1}{3}$

c) The probability the first drawn ball is red and the second is blue =
The probability the first drawn ball is red × The probability the second drawn ball is blue in a condition the first drawn ball is red
$= \frac{4}{10} \times \frac{6}{9} = \frac{4}{15}$

The tree diagram can be used as shown in the figure opposite to find the outcomes of the dependent events.



Try to solve

6) a bag contains 3 red balls and 5 black balls. Two balls are drawn one after another without replacing. What is the probability:

a) the two balls are black?
b) the first ball is black and the second is red?
c) one of the two balls is red and the other is black?

Exercises 2 - 2

1) which of the following events is independent and which is dependent?

a) tossing a coin, then rolling a die once.
b) a card is drawn from a bag without replacing, then another card is drawn from the same bag
c) a ball is drawn from a bag with replacing, then another ball is drawn from the same bag.
d) a football team reaches the semi-final match. If it wins, it will play the championship match.
e) choosing a name by lot without replacing, then choosing another name.
f) choosing a ball from a bag and placing it in another place, then choosing another ball from the Same bag.
g) karem participated and succeeded in the cultural competition on Monday and participated and succeeded in the scientific competition on Thursday, too.

Choose the correct answer:

2) If A and B are two independent events, $P(A) = 0.2$, $P(B) = 0.6$, then $P(A \cup B) =$

a) 0.12   b) 0.32   c) 0.68   d) 0.8

3) If A and B are two independent events $P(A) = 0.25$, $P(B) = 0.4$, then $P(A - B) =$

a) 0.1   b) 0.15   c) 0.3   d) 0.65

4. If A and B are two independent events $P(A) = 0.3$, $P(B) = x$, $P(A \cup B) = 0.72$ then x equals:

   a) 0.24 b) 0.28 c) 0.4 d) 0.6

Answer the following questions:

5. What is the probability of appearing Head and number 3 if a coin is tossed then a die is rolled once?

6. What is the probability of getting Tail 4 times if a coin is tossed four consecutive times?

7. A regular die is rolled once. And A is the event of appearing an even number and B is the event of appearing a squared number. Are A and B two independent events? Explain.

8. If A and B are two events of a sample space of a random experiment and $P(B) = 0.3$, $P(A \cup B) = 0.5$ find the value of $P(A)$ if a and b are:

   a) Two mutually exclusive events. b) Two independent events .

9. A bag contains a set of marbles distributed as follows: 2 red, 3 green and one blue. A marble is randomly chosen with replacing, then another marble is chosen. Find the probability the two marbles chosen are green.

10. In the previous question, if a marble is randomly chosen without replacing, then another marble is chosen, find the probability the first marble chosen is blue and the second is green.

11. A bag contains the following balls: 6 red, 4 orange, 3 yellow, 2 blue and 5 green. A ball is randomly chosen without replacing, then another ball is chosen.
    Find the probability the balls chosen are:

    a) Red and blue. b) Red and yellow. c) Red and red. d) Orange and blue.

12. Two soldiers A and B have shot a bullet in the direction of a target. If the probability the first soldier shot the target is 0.4 and the probability the second soldier shot the same target is 0.7,
    First: find the probabilities of the following events::

    a) The two soldier shot the target together. b) At least one soldier shot the target.
    c) Only one soldier shot the target.. d) At most one of them Shot the target.

    Second: If you know that one soldier at least shot the target, find the probability the soldier A only shot the target.

13. If A and B are two independent events, prove that all the following event pairs are also independent.

    a) A', B' b) A', B c) A, B'

Unit summary

1 *calculating the probability of an event (say A)*

$$P(A) \text{ (the probability of occurring even A)} = \frac{\textbf{the number of outcomes in event n(A)}}{\textbf{the number of outcomes in sample n(s)}}$$

2 *Mutually exclusive events:* They are the two events which do not share in any element and their intersection is the null set ϕ .

If A and B are two mutually exclusive events, then $A \cap B = \phi$

$\therefore$ $P(A \cap B)$ =Zero and $P(A \cup B) = P(A) + P(B)$

3 *Events are not mutually exclusive:* They are the two events in which if an event occurs, it does not trap the occurrence of the other(there are common elements between them) then, $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

4 *Conditional probability:* If S is the sample space of a random experiment and A and B are two events of this space, then, the probability of occurring the event A in condition of occurring the event B and it is denoted by the symbol P(A|B) and read as the probability of occurring event A in condition of occurring event B. it can be determined by the next relation:

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)} \quad \text{where } P(B) > 0$$

5 *the two independent events :* It is said that A and B are two independent events if and only if $P(A \cap B) = P(A) \times P(B)$.

i.e the probability of occurring two independent events together equals the probability of occurring the first event multiplied by the probability of occurring the second event.

It is noticed that if the two events A , B are independent and $P(B) \neq 0$

then $P(A \mid B) = P(A)$

i.e. the occurrence of an event doesn't affect the occurrence of the other event.

6 *Dependent events*

A and B are two dependent events if $\mathbf{P(A \cap B) \neq P(A) \times P(B)}$

because we know from the conditional probability that:

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)} \quad \text{In condition } P(B) \neq 0,$$

$$P(B \mid A) = \frac{P(A \cap B)}{P(A)} \text{In condition } P(A) \neq 0$$

I.e. it can be written as: $P(A \cap B) = P(A \mid B) \times P(B)$

$= P(B \mid A) \times P(A) \qquad P(A) \neq 0, P(B) \neq 0$

In other words, the two events A and B are dependent if the probability of the occurrence of one of them affects in a way the probability of the occurrence of the other event.

General exercises

Choose the correct answer:

1. If A and B are two mutually exclusive events, P(A)= 0.2 and P(B`)= 0.6, then $P(A \cup B) =$

   a 0.4　b 0.6　c 0.8　d 1.2

2. If $A \subset B$, $P(A') = \frac{7}{10}$ and $P(B) = \frac{1}{2}$ then $P(A \mid B)$ equals:

   a $\frac{1}{5}$　b $\frac{3}{5}$　c $\frac{4}{5}$　d $\frac{3}{20}$

3. If A and B are two independent events, P(A)= 0.2 and P(B)= 0.5, then $P(A \cap B) =$

   a 0.7　b 0.4　c 0.3　d 0.1

4. If S={A,B,C} and A,B and C are mutually exclusive events where P(A)= 0.25 and P(B)= 0.4, then P(C) =

   a 0.1　b 0.15　c 0.35　d 0.65

5. If A and B are two independent events of S where P(B)= 0.6 and $P(A \cup B) = 0.68$, then P (A) =

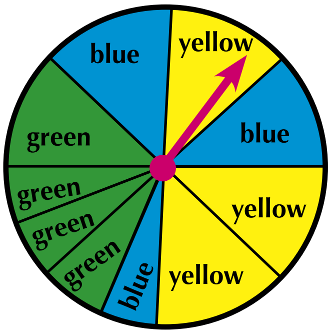
   a 0.2　b 0.3　c 0.4　d 0.5

6. In an experiment to draw a ball randomly from a box containing 10 red balls numbered from 1 to 10 and 5 blue balls numbered by the odd numbers from 1 to 9. A ball has been randomly drawn from the box and it is found that the ball is red. What is the probability the ball drawn is numbered 9?

7. If 25% of the students at one of a school grade failed in the math exam and 10% failed in math and chemistry exams and a student has been randomly chosen. What is the probability the student chosen failed in chemistry if he (she) failed in math?

8. Family: A 3 -kid family. If A: is the event the family has male and female kids and B: is the event the family has one male kid at most. Are the two events A and B independent? Explain.

9. A regular die has been rolled once and the spinning wheel shown opposite has been spined, find the probabilities of appearing the following events:



   a P(the number 3 then the green sector)
   b P(a prime number then the blue sector
   c P(the number 5 then the yellow sector
   d P(an even number then the green sector)

10. If A and B are two mutually exclusive events from a sample space of a random experiment where:
    $P(A - B) = \frac{1}{2}$, $P(A \cup B) = \frac{5}{8}$, Find: P (A) , P (B) , $P(A' \cup B')$

11. a regular die has been rolled two consecutive times, find the probability:

    a Appearing the number 5 on one face known that the same number appeared on both faces.
    b Appearing the number 4 on one face known that the two number appeared on the two faces each is greater than 3.
    c Non-appearing the number 3 on any of the two faces known that the two numbers appeared are odd.

12. In a competition, a question has been given to the two competitors A and B. if the probability the competitor A answer the question is 0.6 and the probability the competitor B answer the same question is 0.8, find the following probabilities:
   a. A and B answer the question together
   b. one of them answer the question at least.
   c. no one answer the question.

13. A bag contains 26 cards; 10 red and 16 green, two cards have been randomly chosen one after another without replacing. What is the probability:
   a. the two cards are red?
   b. the two cards are green?
   c. the first card is red and the second is green?
   d. the first card is green and the second is red?

14. A competition has been conducted to form two teams of the students where cards are randomly drawn out of 9 cards numbered from 1 to 9. If:
   Team A= The students who draw the odd numbered cards
   Team B= The students who draw the even numbered cards.
   a. what is the probability the student draws number 7 if he (she) is from team A?
   b. what is the probability the student draws number 4 if he (she) is from team B?

15. the next table shows the distribution of 50 people in respect to smoking and the diseases related.

b \ a	Sick	Not sick
Smoker	30	10
Nonsmoker	2	8

   If a person is randomly chosen from these people, find each of the following:
   a. the probability this person is sick.
   b. the probability this person is sick in a condition he (she) is smoker.
   c. the probability this person is sick in a condition he (she) is nonsmoker.

16. the next table shows the distribution of 10 people:

	Students	Employee
Athlete	4	1
Not athlete	3	2

   If two persons are randomly chosen, find the probability one of them is athlete and the other is not athlete.
   a. if the choice is done with replacing.
   b. if the choice is done without replacing.

For more activities and exercises, visit www.sec3mathematics.com.eg

Accumulative test

Short-answered questions:

First: complete the following:

1. The probability of occurring the impossible event = ...
2. When a die is rolled once, the probability the number 3 appears on the upper face = ...
3. If a digit of the number 37450 is randomly chosen, the probability the number chosen is even = ...
4. The value for the correlation coefficient if the correlation is perfect direct = ...

Second: choose the correct answer:

5. when a regular die is rolled once and observing the upper face, the probability of appearing a number greater than or equal 6 equals:

   a Zero b $\frac{1}{6}$ c $\frac{5}{6}$ d 1

6. A basket contains 48 balls of the same type; some are white, some others are red and the rest are green. If the probability of drawing a red ball equals $\frac{5}{8}$, then the number of the red balls in the basket equals:

   a 24 b 30 c 32 d 36

7. The weakest correlation coefficient of the following is::

   a -0.9 b -0.5 c 0.1 d 0.4

Long-answered questions:

8. Calculate Spearman's correlation coefficient from the next table:

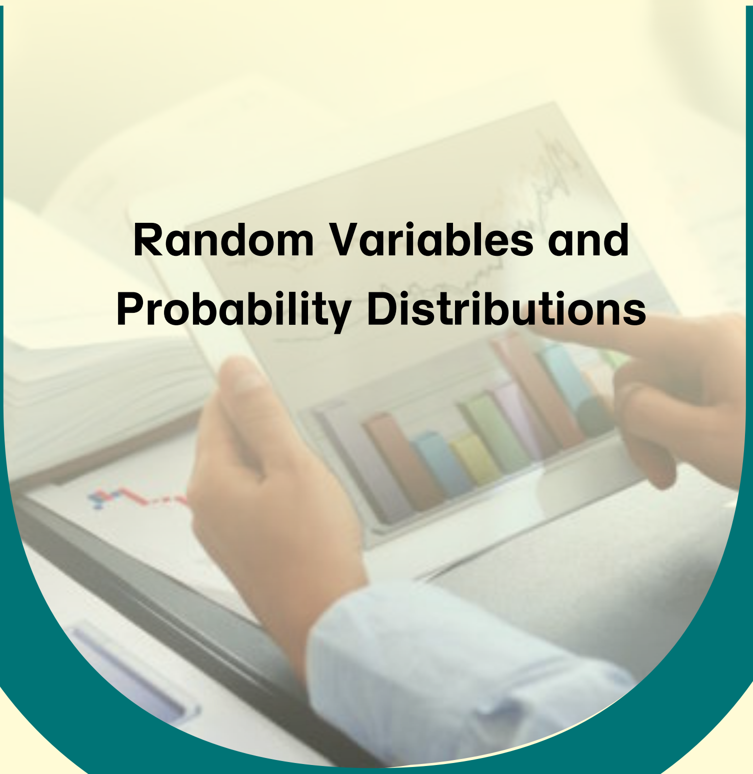
Math marks	Excellent	Good	Very good	Pass	Very good	Pass	good
Physics marks	Weak	Very good	Pass	Very good	Pass	Very good	good

9. In an international conference of 150 members, it is found that 100 members speak English, 60 members speak French and 20 members speak both English and French. A member is randomly chosen, find the probability the member chosen:
   a Speaks one language of both at least.
   b Speaks English if he (she) speaks French.
   c Speaks French if he (she) speaks English

10. A bag contains 12 yellow balls, 8 red balls and two balls are drawn one after another without replacing. What is the probability:
   a The two balls are yellow?
   b The first ball is yellow and the second is red?
   c The two balls are red?

Unit
3

Random Variables and Probability Distributions



introduction

You have previously learned the random experience and some probability concepts.

You sometimes like to deal with numerical values related to the results of the random experience which may be characteristics that are mathematically difficult to deal with. In such a case, we convert these descriptive values into real numerical values called a random variable, which is used to express the results of the random experience.

In this unit, we are going to learn two types of random variables which are:

- Discrete Random Variables
- Continuous Random Variables

Furthermore, we will learn the probability distribution functions of the random variables which are divided into:

- Probability distribution function of discrete random variable.
- Probability density function.

Unit objectives

By the end of the unit and carrying the involved activities, the students should be able to:

- Identify the concept of the random variable and distinguish the discrete random variable and the continuous one.
- Identify the probability density function of a continuous random variable and use it to calculate the probability of occurring the value of the random variable within a certain interval.
- Identify the concept of the mean (expectation) and variance.
- Deduce the standard deviation of a random variable.
- Identify the coefficient of variation
- Identify the continuous distributions.

Key Terms

- *Random Variable*
- *Discrete Random Variable*
- *Probability Distributions*
- *Expectation(Mean)*
- *Variance*
- *Coefficient of Variation*
- *Probability Density*

Lessons of the unit

Lesson (1 - 1): Discrete random variable.

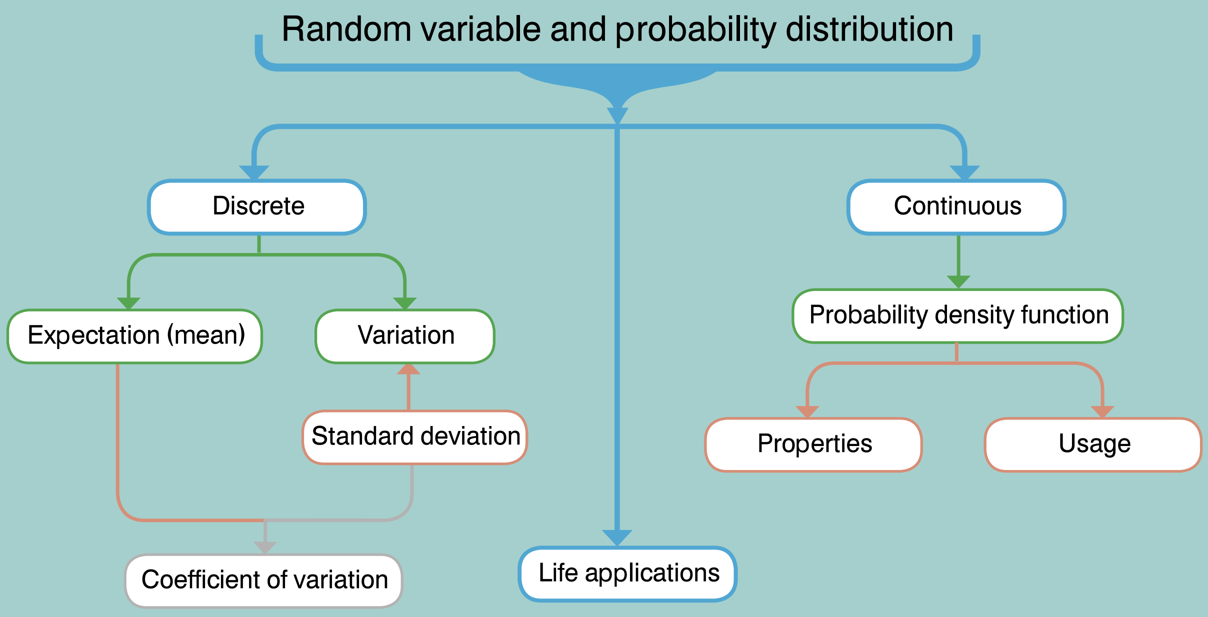
Lesson (1 - 2): Expectation (mean) and variance of the discrete random variable.

Lesson (1 - 3): The probability density function of the continuous random variable

Materials

Scientific Calculator

Unit planning organizer



Unit Three

3 - 1 Discrete random variable

You will learn

- Random variable
- Discrete random variable
- Continuous random variable
- Probability distribution

Key terms

- Random Variable
- Discrete Random Variable
- Continuous Random Variable
- Probability Distributions

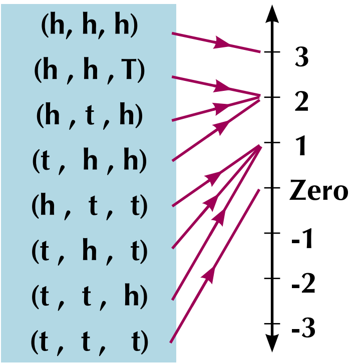
Introduction: : you have learned the random experiment and found the sample space for this experiment . In this lesson, you will identify a new variable related to such an experiment - The random variable .

You will learn in this lesson how to describe the items of two different phenomena with respect to the relation between them.

Random variable:

In an experiment for tossing a coin three consecutive times, the sample spaces S can be identified as shown in the opposite figure. If you are asked to find " the number of images" which appear in the sample space S, you draw a diagram to show the relation between S (as an independent variable) and the number of the image which is real number R (as a dependent variable).

This relation expresses a function and can be symbolically written as : X: S → R where X is denoted the random variable .



Definition

The random variable is a function whose domain is the set of the elements of the sample space S and Co-domain is the set of real numbers R.

Remember

The function is defined by:

- **Domain**
- **Co-domain**
- **Base of the function**

The function range is the set of the images of the domain elements in the co-domain.

Then, the random variable range X in the previous example = {0 , 1 , 2 , 3 }

Note: The random variable divides the sample space S into mutual exclusive events; each event is linked with a real number , this link expresses a function X from the sample space S to the set of the real number R.

Discrete Random Variable

Definition

The range of the discrete random variable is a finite set (i.e. it can be counted) of the real numbers.

Materials

- Scientific Calculator , Computer graphics.

For examples :

- the number of stocks specialized for an individual in an underwriting of a joint-stock company.
- The number of car accidents on a high way within a week.
- The number of the outcoming telephone calls for a family within a month.

Example **Discrete random variable**

1 In an experiment for tossing a coin three consecutive times , If the random variable X expresses " number of heads - number of tails" , write down the random variable range.

Solution

S = { (H , H , H) , (H , H , T) , (H , T , H) , (H , T , T) , (T , H , H) ,
(T , H , T) , (T , T , H) , (T , T , T) }

Sample space S	X: number of tails – number of heads
(H , H , H)	3 – 0 = 3
(H , H , T)	2 – 1 = 1
(H , T , H)	2 – 1 = 1
(H , T , T)	1 – 2 = – 1
(T , H , H)	2 – 1 = 1
(T , H , T)	1 – 2 = – 1
(T , T , H)	1 – 2 = – 1
(T , T , t)	0 – 3 = – 3

The random variable range = {- 3 , - 1 , 1 , 3 }

Try to solve

1 in the previous example, find the random variable range expressing : number of heads × number of tails.

Example **Discrete random variable**

2 A regular die has been tossed two consecutive times . Find the random variable expressing the sum of the two shown numbers .

Solution

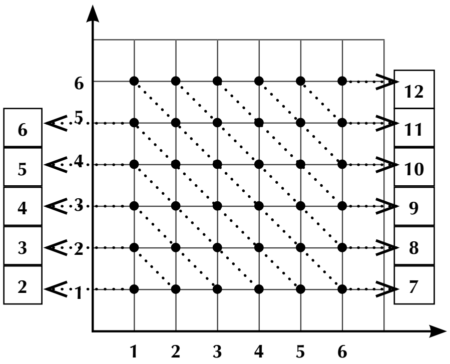
Sample space S	x: sum of the two numbers
(1 , 1)	2
(1 , 2) , (2 , 1)	3
(3 , 1), (2 , 2), (1 , 3)	4
(4 , 1) , (3 , 2) , (2 , 3) , (1 , 4)	5
(5 , 1), (4 , 2), (3 , 3), (2 , 4), (1, 5)	6

Sample space S	x: sum of the two numbers
(4 , 3) ,(3 , 4) ,(2 , 5) ,(1 , 6), (6 , 1) , (5 , 2)	7
(6 , 2), (5 , 3), (4 , 4), (3 , 5), (2 , 6)	8
(6 , 3), (5 , 4), (4 , 5), (3 , 6)	9
(6 , 4) , (5 , 5) , (4 , 6)	10
(6 , 5) , (5 , 6)	11
(6 , 6)	12

From the previous table , the random variable range is

x = {2 , 3 , 4 , 5 , 6 , 7 , 8 , 9 , 10 , 11 , 12 }

we can use the lateral shape to find the random variable range x .



Try to solve

2 In the previous example, find the random variable range expressing : (the greater of the two shown numbers).

Probability distribution

Probability Distribution Function of Discrete Random Variable

Definition

If x is a discrete random variable whose range is the set:{ x_1 , x_2 , x_3. ...... , x_r } then the function f defined as follows: $f(x_r) = P(X = x_r)$ for each r = 1 , 2 , 3 , ...

Identifies what is called the probability distribution function of the discrete random variable X which is expressed by a set of ordered pairs specifying the function F.

i.e. the probability distribution of the random variable X = { (x_1 , $f(x_1)$) , (x_2 , $f(x_2)$) , (x_3 , $f(x_3)$) , ... , (x_n , $f(x_n)$) }

Note : the probability distribution of the random variable x can be written in the form of the following table:

x_r	x_1	x_2	x_3	......	x_n
$f(x_r)$	$f(x_1)$	$f(x_2)$	$f(x_3)$	......	$f(x_n)$

You can notice that the function f in the previous definition satisfies the following two conditions:

1- $f(x_r) \geqslant 0$ for each **r** = 1 , 2 , 3 , ... , n

2- $f(x_1) + f(x_2) + f(x_3) + ... + f(x_r) = 1$

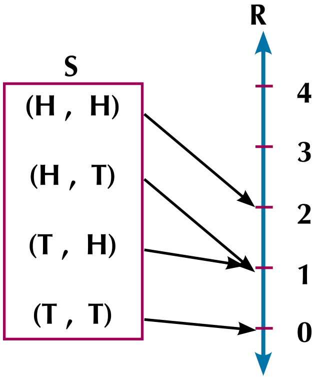
Example probability distribution function

3 A coin has been tossed three consecutive times and the upper face has been seen. Write down the probability distribution function of the discrete random variable x expressing the number of appearing the head face.

Solution

S = { (H , H) , (H , T) , (T , H) , (T , T) }

From the lateral shape , we find that the random variable range expressing the number of appearing the head = { 0 , 1 , 2 }



$f(0) = p(X = 0) = \frac{n(x_1)}{n(s)} = \frac{1}{4}$

$f(1) = p(X = 1) = \frac{n(x_2)}{n(s)} = \frac{2}{4}$, $f(2) = p(X = 2) = \frac{n(x_3)}{n(s)} = \frac{1}{4}$

The probability distribution function is:

x_r	0	1	2
$f(x_r)$	$\frac{1}{4}$	$\frac{2}{4}$	$\frac{1}{4}$

Try to solve

3 In the previous example, write down the probability distribution function of the discrete random variable x expressing:(number of appearance of head - number of appearance of tail).

Example **Drawing without replacing**

4 A box contains 5 identical cards numbered from 1 to 5. Two cards have been drawn one after another without replacing , find the probability distribution function for the random variable expressing the least number of the two numbers on the two drawn cards..

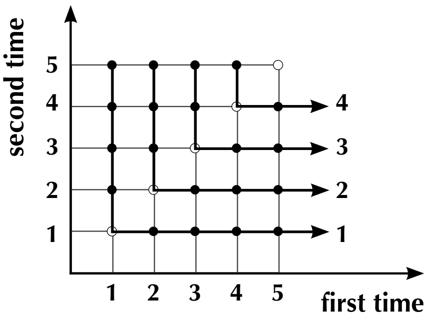
Solution

As long as the cards have been drawn without replacing to the box, the cards drawn is not repeated more, in other words, the pairs of cards carrying the digits (1 , 1) , (2 , 2) , (3 , 3) , (4 , 4) , (5 , 5) are not included in the sample space as shown in the opposite figure.

$n(s) = 20$

From the figure opposite, the range of random variable x is:

$x = \{ 1 , 2 , 3 , 4 \}$ and:



$f(1) = p(X = 1) = \frac{8}{20}$

$f(2) = p(X = 2) = \frac{6}{20}$

$f(3) = p(X = 3) = \frac{4}{20}$

$f(4) = p(X = 4) = \frac{2}{20}$

the probability distribution function of the discrete random variable x is given as shown in the following table :

x_r	1	2	3	4
$f(x_r)$	$\frac{8}{20}$	$\frac{6}{20}$	$\frac{4}{20}$	$\frac{2}{20}$

Try to solve

4 In an experiment of rolling a regular die two consecutive times and observing the seen face in each time, find the probability distribution function of the discrete random variable expressing the greatest number of the two numbers appeared on the two upper faces.

Example

Using the rule (base) of the function

(5) If X is a discrete random variable and its probability distribution function is determined by the relation :

$f(x) = \frac{k + 2x}{24}$ where x = 0 , 1 , 2 , 3 find the value of k, then write down the probability distribution function.

Solution

$\because f(0) = p(X = 0) = \frac{k}{24}$, $f(1) = p(x = 1) = \frac{k+2}{24}$,

$f(2) = p(X = 2) = \frac{k+4}{24}$, $f(3) = p(x = 3) = \frac{k+6}{24}$

$\because p(X = 0) + p(X = 1) + p(X = 2) + p(X = 3) = 1$

$\therefore \frac{k}{24} + \frac{k+2}{24} + \frac{k+4}{24} + \frac{k+6}{24} = 1$

$\therefore \frac{k + k + 2 + k + 4 + k + 6}{24} = 1$ $\therefore 4k + 12 = 24$

$\therefore 4k = 24 - 12$ $\therefore 4k = 12$ $\therefore k = 3$

To find the probability distribution function , you should find:

$p(X = 0) = \frac{k}{24} = \frac{3}{24}$, $p(X = 1) = \frac{k+2}{24} = \frac{5}{24}$

$p(X = 2) = \frac{k+4}{24} = \frac{7}{24}$, $p(X = 3) = \frac{k+6}{24} = \frac{9}{24}$

$\therefore$ The probability distribution function is:

x_r	0	1	2	3
$f(x_r)$	$\frac{3}{24}$	$\frac{5}{24}$	$\frac{7}{24}$	$\frac{9}{24}$

Try to solve

(5) If X is a discrete random variable whose range = { 1 , 2 , 3 } and its probability distribution function is determined by the relation $f(x) = \frac{ax}{9}$. Find the value of a , then write down the probability distribution function.

Exercises 3 - 1

First : choose the correct answer:

1) Which function of the following represents a probability distribution function of the discrete random variable X:

a

x_r	1	2	3	4
$f(x_r)$	0.06	0.15	0.42	0.26

b

x_r	0	1	3	5
$f(x_r)$	0.5	0.3	0.4	−0.2

c

x_r	−2	−1	1	2
$f(x_r)$	0.32	0.14	0.23	0.31

d

x_r	3	4	5	6
$f(x_r)$	0.23	0.32	0.17	0.18

2) If X is a random variable whose range is {0, 1, 2} then all the following functions do not represent its probability distribution function except :

a $f(x) = \frac{x^2 + 1}{8}$ b $f(x) = \frac{2x + 1}{3}$ c $f(x) = \frac{1}{x + 2}$ d $f(x) = \frac{3x - 1}{6}$

3) If X is a random variable whose range is {1 , 2 , 3}, $p(x = 1) = 0.3$ and $p(X = 2) = 0.5$ then $p(X = 3)$ is equal:

a 0.1 b 0.2 c 0.7 d 0.8

4) If x is a random variable whose range is {1 , 2 , - 1 , 0 }, $p(X = -1) = 0.2$ and $p(X = 0) = 0.4$, $p(X = 1) = 0.1$ then $p(X > 1)$ equal :

a 0.3 b 0.4 c 0.5 d 0.6

5) in an experiment for tossing a coin three consecutive times and X is the discrete random variable expressing (number of heads - number of tails) , then the range of X is :

a {1 , 3} b {0 , 1 , 3} c { 0 , 1 , 2 , 3} d {-3 , -1 , 1 , 3}

6) If X is a discrete random variable whose range = { 0, 1, 2} and its probability distribution function is determined by the relation $f(x) = \frac{ax}{6}$ then the value of a is equal:

a $\frac{1}{2}$ b 1 c $\frac{3}{2}$ d 2

Second : Answer the following questions:

7) The two following tables show the probability distribution of a discrete random variable x. Find the value of a in each table:

a

x_r	1	2	3
$f(x_r)$	a	2a	3a

b

x_r	-2	−1	0	1	2
$f(x_r)$	a	0.2	0.3	3a	a

c

x_r	0	1	3	4
$f(x_r)$	a	$2a^2$	$3a^2$	3a

8. If X is a discrete random variable whose range = { 0 , 1 , 2 , 3 } and the values of p (X = 0) = 0.2 , p (X = 1) = 0.33 and p (X = 2) = 0.37 find the probability distribution of the discrete random value X.

9. If the values of the discrete random variable X in a random experiment are: -2 , 0 , 2 , 4 with probabilities of $\frac{m}{5}$, $\frac{m+1}{5}$, $\frac{2m-1}{5}$ and $\frac{3m-2}{5}$ respectively , find the value of m , then write down the probability distribution function of the variable X.

10. If X is a discrete random variable whose probability distribution function is determined by the relation : $f(x) = \frac{2a + 3x}{54}$ and the range of x = { 1 , 2 , 3 , 4 }find the value of a and write down the probability distribution function of the variable X.

11. If X is a discrete random variable whose probability distribution is determined by the function $f(x) = \frac{k + 3x}{50}$ where x = 1 , 2 , 3, 4, find the value of K, then write down the probability distribution of the variable X.

12. In an experiment for tossing a coin three consecutive times, if the discrete random variable X expresses " number of heads- number of tails" , write down the probability distribution of the variable X.

13. Two boxes each contains three numbered balls from 3 to 5. A ball has been randomly drawn from each box and the discrete random variable x is defined as " the sum of two numbers" existed on the two drawn balls. Find the probability distribution of the discrete random variable X.

14. In an experiment for tossing a die two consecutive times and observing the number on the upper face in each time , write down the probability distribution of the discrete random variable X expressing " the least number of the two observed numbers".

15. A box contains 4 balls numbered from 1 to 4, two balls have been drawn one after another (with replacing), write down the probability distribution of the discrete random variable x expressing " the mean of the two numbers on the two drawn balls".

16. If X is a discrete random variable expressing the number of girls in a three- kid family, write down the range of the discrete random variable X . If we suppose that the probability of giving a birth of a boy is equal to the probability of giving a birth of a girl and disregarding having twins, find the probability distribution of the random variable X (the order of boys and girls is taken into acount)

Expectation (mean) and Variance of a Discrete Random Variable

Unit Three

3 - 2

You will learn

- Expectation (Mean)
- Variance
- Standard deviation
- Coefficient of variation

Key terms

- Expectation (Mean)
- Variance
- Coefficient of Variation

Introduction : To identify the characteristics of the probability distribution (identifying the characteristics of the original community or comparing different communities) , the basic features are necessarily needed to measure its mean value which the possible values of the discrete random value accumulate around it and is called expectation (mean), There are also other values measuring the dispersion of the discrete random variable values from the value of the mean and is known as variance, Therefore, expectation and variance summarize the most important characteristics of the discrete random variable.

Expectation (Mean)

Expectation is the value at which the most values of the discrete random variable are centralized and sometimes is called " mean" and is denoted by the symbol (μ) and read as (Mu).
If X is a discrete random variable whose probability distribution function is f and its range is: $\{ x_1 , x_2 , x_3 , ... , x_n\}$ by probabilities of $f(x_1) , f(x_2) , f(x_3) , ... , f(x_n)$ respectively, then the expectation is given by the relation:

$$\text{Expectation } (\mu) = \sum_{r=1}^{n} x_r \times f(x_r)$$

i.e.: Expectation $(\mu) = x_1 \times f(x_1) + x_2 \times f(x_2) + x_3 \times f(x_3) + ... + x_n \times f(x_n)$

Example

1 If x is a discrete random variable whose probability distribution is shown in the below table:

x_r	-1	0	1	2	3
$f(x_r)$	0.3	0.1	0.1	**a**	0.2

First : Find the value of a **Second :** Find the expectation (Mean)

Solution

First : you know that the sum of probabilities is equal one

$\because p(X = -1) + p(X = 0) + p(X = 1) + p(X = 2) + p(X = 3) = 1$

$\therefore 0.3 + 0.1 + 0.1 + \mathbf{a} + 0.2 = 1$

$\therefore \mathbf{a} + 0.7 = 1 \qquad \therefore \mathbf{a} = 1 - 0.7 = 0.3$

Materials

- Scientific calculator - Computer graph programs.

Second :

$\because$ Expectation $(\mu) = \sum_{r=1}^{n} x_r \times f(x_r) = -1 \times 0.3 + 0 \times 0.1 + 1 \times 0.1 + 2 \times 0.3 + 3 \times 0.2$

$= -0.3 + 0 + 0.1 + 0.6 + 0.6 = 1$

Try to solve

1. If X is a discrete random variable whose range = {0 , 1 , 2 , 3 , 4} , $p(X = 0) = p(X = 4) = \frac{1}{16}$, $p(X = 1) = p(X = 3) = \frac{1}{4}$
Find: **first:** p (X = 2) **second:** Expectation

Example

2. If x is a discrete random variable whose probability distribution is shown as follows:

x_r	0	1	2	b	6
$f(x_r)$	0.1	0.1	0.3	**a**	0.3

Find the value of a and b if the expectation $\mu = 3.5$

Solution

From the properties of the probability distribution: f(0) + f(1) + f(2) + f(b) + f(6) = 1

$\therefore 0.1 + 0.1 + 0.3 + \mathbf{a} + 0.3 = 1$ $\therefore \mathbf{a} = 1 - 0.8$ **a = 0.2**

$\because$ expectation $(\mu) = \sum_{r=1}^{n} x_r \times f(x_r) = 3.5$

$\therefore 0 \times 0.1 + 1 \times 0.1 + 2 \times 0.3 + b \times 0.2 + 6 \times 0.3 = 3.5$

$\therefore 0 + 0.1 + 0.6 + 0.2\,b + 1.8 = 3.5$ $\therefore 0.2\,b = 3.5 - 2.5$

$\therefore b = 1 \div 0.2 = 5$ **b = 5**

Try to solve

2. If X is a discrete random variable whose probability distribution is shown as follows:

x_r	0	2	3	4
$f(x_r)$	$\frac{3}{16}$	2 p	$\frac{1}{16}$	p

First : find the value of p. **Second :** find the expectation

Variance

The variance of a discrete random variable measures the amount of dispersion of the discrete random variable of its expected value and is denoted by the symbol (σ^2) and read as (sigma squared) and is given by the relation:

$$\sigma^2 = \sum_{r=1}^{n} x_r^2 \times f(x_r) - \mu^2$$

Note: The standard deviation of a discrete random variable x is the square root of the variance and is denoted by the symbol σ . It's noticed that the variance and the standard deviation are always positive quantities .

Example

3 If X is a discrete random variable whose probability distribution function is $f(x) = \frac{x+4}{16}$ where x = -2 , m , 1 , 2 find the value of m, then find the mean and the variance of the discrete random variable X.

Solution

From the properties of the probability distribution:

$\because p(X = -2) + p(X = m) + p(X = 1) + p(X = 2) = 1$

$\therefore \frac{2}{16} + \frac{m+4}{16} + \frac{5}{16} + \frac{6}{16} = 1$

$\therefore \frac{17+m}{16} = 1$ $\quad\quad \therefore 17 + m = 16 \quad\quad \therefore m = -1$

x_r	$f(x_r)$	$x_r \cdot f(x_r)$	$x^2_r \cdot f(x_r)$
-2	$\frac{2}{16}$	$\frac{-4}{16}$	$\frac{8}{16}$
-1	$\frac{3}{16}$	$\frac{-3}{16}$	$\frac{3}{16}$
1	$\frac{5}{16}$	$\frac{5}{16}$	$\frac{5}{16}$
2	$\frac{6}{16}$	$\frac{12}{16}$	$\frac{24}{16}$
		$\frac{5}{8}$	$\frac{5}{2}$

The expectation $(\mu) = \sum_{r=1}^{n} x_r \times f(x_r) = \frac{5}{8}$

The variance $(\sigma^2) = \sum_{r=1}^{n} x_r^2 \times f(x_r) - \mu^2 = \frac{5}{2} - (\frac{5}{8})^2 = \frac{135}{64}$

Try to solve

3 If X is a discrete random variable whose probability distribution is defined by the function where find $f(x) = \frac{a}{x+1}$ where x = 0 , 1 , 2 , 3 find:

First: the value of a

Second: The expectation and the standard deviation of the discrete random variable x.

Coefficient of Variation

As you were learning the standard deviation as a measure for the dispersion of the values of the discrete random variable of its expectation, you knew that it is measured by the same measuring units of the variable investigated whether these units are in degrees, meters or kilograms... i.e. it is proper to compare two sets of the same measuring units and the same means. If the measuring units or means are different between the two sets, the standard deviation is no longer proper to be used as a measure for the comparison, As a result, the need of a relative measure for dispersion can help us get rid of such different units. The coefficient of variation represents a proper solution to such a problem.

The coefficient of variation for any set of items is known as the percentage between the standard deviation of the set and its expectation (mean) and it can be defined as in the following relation :

$$\textbf{Coefficient of variation} = \frac{\textbf{standard deviation}}{\textbf{mean}} \times 100\ \% = \frac{\sigma}{\mu} \times 100\%$$

This coefficient forms the dispersion of the set in the form of a percentage abstracted from distinction where the measured units are not affected by the phenomenon.

Example

4 If the expectation and the standard deviation for the marks of a group of students in history and Geography are as follows , known that the full mark is 100 marks.

Subject	history exam	geography exam
Expectation	70	96
Standard deviation	7	8

Find the coefficient of variation for each subject - What do you notice?

Solution

$\because$ Coefficient of variation $= \frac{\text{standard deviation}}{\text{mean}} \times 100\ \%$

$\therefore$ The coefficient of variation of history $= \frac{7}{70} \times 100\ \% = 10\ \%$.

The coefficient of variation of geography $= \frac{8}{96} \times 100\ \% \simeq 8.3\ \%$

From the Solution: we observe that the relative dispersion of the history exam is greater than of geography. This means the geography exam is more homogeneous than the history exam.

Try to solve

4 If a factory produces two types of electric lamps A, B and their average of use is 1850, 1580 hours and the standard deviation of both is 250, 230 respectively, find the coefficient of variation of each type. What do you notice.

Example

5 A box contains 6 cards; 2 cards carry the number two, 3 cards carry the number 3 and a card carries the number 11, if a card has been randomly drawn and the discrete random variable X is defined as " the appeared number on the drawn card"

Find:

a The probability distribution function of the variable X.

b Expectation and standard deviation of the variable X **c** the coefficient of variation.

Solution

a X takes the values 2 , 3 , 11 where: $f(2) = p\ (X = 2) = \frac{2}{6} = \frac{1}{3}$

$f(3) = p\ (X = 3) = \frac{3}{6} = \frac{1}{2}$, $f\ (11) = p(X = 11) = \frac{1}{6}$

the table below illustrated the probability distribution function of the discrete random variable x.

x_r	2	3	11
$f(x_r)$	$\frac{1}{3}$	$\frac{1}{2}$	$\frac{1}{6}$

To calculate the expectation and the standard deviation, you ought to form the table below:

x_r	$f(x_r)$	$x_r \cdot f(x_r)$	$x^2_r \cdot f(x_r)$
2	$\frac{2}{6}$	$\frac{4}{6}$	$\frac{8}{6}$
3	$\frac{3}{6}$	$\frac{9}{6}$	$\frac{27}{6}$
11	$\frac{1}{6}$	$\frac{11}{6}$	$\frac{121}{6}$
Total		4	26

b Expectation $(\mu) = \sum_{r=1}^{n} x_r \times f(x_r) = 4$

Variance $(\sigma^2) = \sum_{r=1}^{n} x^2 f(x_r) - \mu^2 = 26 - (4)^2 = 10$

standard deviation $\sigma = \sqrt{10} = 3.16$

c $\because$ Coefficient of variation $= \frac{\text{standard deviation}}{\text{mean}} \times 100\ \%$

$\therefore$ Coefficient of variation $= \frac{3.16}{4} \times 100\ \% = 79\ \%$

Try to solve

5 A box contains 10 cards ; a card carries the number 1 , two cards each carries the number 2, 3 cards each carries the number 3, 4 cards each carries the number 4. If a card is randomly drawn and the discrete random variable X expresses the number on the drawn card, find the probability distribution function of this variable and calculate each of the expectation , its standard deviation and the coefficient of variation.

Exercises 3 - 2

First : Choose the correct answer:

1. If the probability distribution of the discrete random variable X is {(0 , 0.25) , (1 , 0.5) , (2 , 0.25) } then its expectation equals:
   a 0.5 b 1 c 1.25 d 1.5

2. If x is a discrete random variable and the expectation equals 0.6 , $\sum_{r=1}^{n} x_r^2 \times f(x_r) = 4.36$ then its standard deviation equals:
   a 1.94 b 2 c 3.76 d 4

3. If x is a discrete random variable and the expectation equals 0.4 , $\sum_{r=1}^{n} x_r^2 \times f(x_r) = 6.16$ then its variance equals :
   a 2.4 b 5.76 c 6 d 6.56

Second : Find the expectation and the standard deviation for the probability distribution for each of the following:

4.

x_r	2	3	9
$f(x_r)$	$\frac{1}{3}$	$\frac{1}{2}$	$\frac{1}{6}$

5.

x_r	–5	–4	1	2
$f(x_r)$	$\frac{1}{24}$	$\frac{3}{8}$	$\frac{5}{12}$	$\frac{1}{6}$

6.

x_r	–3	–1	0	1	2	3
$f(x_r)$	$\frac{1}{12}$	$\frac{1}{6}$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{6}$	$\frac{1}{12}$

Third : Answer the following questions:

7. If X is a discrete random variable whose probability distribution is shown in the table below:

x_r	1	2	4	6
$f(x_r)$	0.2	0.3	**a**	0.1

First : find the value of a. Second : find the mean and the standard deviation

8. If the range of the discrete random variable X is { 1 , 2 , 3 , 4 }, $p(X = 1) = \frac{4}{25}$, $p(X = 2) = \frac{7}{25}$ and $p(X = 4) = \frac{1}{5}$, find the expectation and variance of X .

9. If X is a discrete random variable whose range { 0 , 1 , 2 , 3 , 4 } , $p(X = 0) = p(X = 4) = \frac{1}{16}$ and $p(X = 1) = p(X = 3) = \frac{1}{4}$ Find: first : p (X : 2) second: the mean and the variance of the variable X .

10. If X is a discrete random variable whose probability distribution function is shown in the table below where $0 < h < 1$

x_r	–3	zero	3	6
$f(x_r)$	h	h^2	$2h^2$	h

Find:

a the value of h

b The probability distribution of the variable X.

c The mean and variance of the variable X

11 If x is a discrete random variable whose probability distribution is shown in the table below:

x_r	1	2	4	a
$f(x_r)$	0.2	0.3	0.4	0.1

Calculate the value of a if the expectation $\mu = 3$, then find the standard variation of the discrete random variable X .

12 If the probability distribution of a discrete random variable X is defined by the function f where : $f(x) = \frac{a\,x}{9}$, where x = 1, 2, 3 . find:

a the value of a   b calculate the expectation and the variance of the variable X.

13 If X is a discrete random variable whose probability distribution is defined by the function: $f(x) = \frac{x^2 + 1}{a}$ where x = 0, 1 , 2 , 3 find: a the value of a b calculate the coefficient of variation of the variable x.

14 If X is a discrete random variable whose probability distribution is defined by the function: $f(x) = \frac{x + 4}{16}$ where x = - 2, m, 1, 2 find: a the value of m b the mean and variance of the variable x.

15 If x is a discrete random variable whose probability distribution is defined by the function f where: $f(x) = \frac{a}{x + 3}$, x = 0 , 1 , 2 , 3

a find the value of a   b the expectation and variance.

16 If the range of the discrete random variable X is { -1 , 0 , 2 } , $p(X = -1) = \frac{1}{4}$ and the expectation equals 1 , find:

a p(X = 0) , p (X = 2)   b The coefficient of variation.

17 If x is a discrete random variable whose mean $\mu = 3$ and its probability distribution is as follows :

x_r	0	2	k	4
$f(x_r)$	1	2a	$\frac{1}{4}$	5a

a Calculate the values of a and k

b Find the standard deviation of the variable X.

Unit 3

3 - 3 Probability density function of the continuous random variable

You will learn

- Probability density function

Key terms

- *Probability Density*

Continuous Random Variable

Definition

The continuous random variable : its range is an interval of the real numbers (closed or opened) . i.e. it's a non - counted set of real numbers .

For examples:

- The wage of a government worker randomly chosen.
- The temperature expected on a day.
- The length of a basketball candidate.

Example — Continuous random variable

6. The point (x , y) is located inside or on the circle $x^2 + y^2 = 4$ whose center is the origin O and its radius length is 2 units. Required is to find the range of the random variable X expressing how distant the point from the center of the circle

Solution

$\because f = \{ (x , y): x^2 + y^2 \leqslant 4\}$

$\therefore 0 \leqslant a \leqslant 2$ where a is the distance between from point (x , y) to the center of the circle.

$\therefore$ the range of the random variable x = [0 , 2]

we can notice that each point in this interval is a possible value to the random variable X as shown in the figure

Try to solve

1. If the maximal life span of a type of cellular phones (X) is approximated by 18 working hours, write down the range of X.

Try to solve

2. Show which of the following represents a discrete random variable and which represents a continuous random variable.

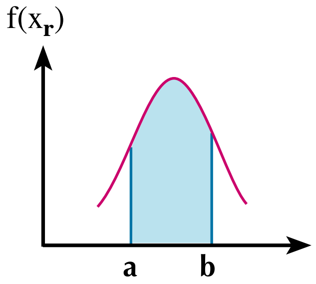
   a. Number of the bread leaves produced by a bakery within an hour.
   b. The time kareem takes to wait his friend Ziad.
   c. Number of the goals scored by the team won the handball match.
   d. The number of the traffic violations recorded on Cairo - Alex desert rood on a day.
   e. The time taken by a teacher to explain the lesson of the random variable.

Materials

- Scientific calculator, computer graph programs.

Probability Density Function

For any continuous random variable x, there is a real function whose range is non- negative, and denoted by the symbol f(x) called the probability density function by which the probabilities of events expressing it by the random variable can be found throughout the area included beneath the function curve and on the x axis and p (a<x<b) can be calculated by finding the area of the shaded part of the curve of the function f between the two values a and b as shown in the opposite figure .



This function satisfies the next conditions:

- $f(x) \geqslant 0$ for all the values of x belonging to the domain of the function.
- The area of the region located beneath the curve of the function f and on the x axis equal 1.

Example

1 If x is a continuous random variable whose probability density function is:

$$f(x) = \begin{cases} \frac{1}{6}(2x - 1), & 1 \leqslant x \leqslant 3 \\ \text{zero}, & \text{otherwise} \end{cases}$$

a **Prove that :** $p(1 < X < 3) = 1$

b **Find :** $p(X \leqslant 2)$, $p(X > 2.5)$, $p(2 \leqslant X \leqslant 2.5)$.

Solution

Remember

Area of the rectangle = length × width

Area of the triangle = $\frac{1}{2}$ base length × height

Area of the trapezoid = $\frac{1}{2}$ the sum of the two parallel bases × height

$f(1) = \frac{1}{6} \times (2 - 1) = \frac{1}{6}$

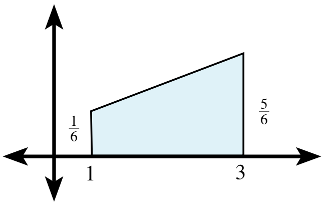
$f(3) = \frac{1}{6} \times (6 - 1) = \frac{5}{6}$

$f(2) = \frac{1}{6} \times (4 - 1) = \frac{3}{6}$

$f(2.5) = \frac{1}{6} \times (5 - 1) = \frac{4}{6}$

a $p(1 \leqslant x \leqslant 3) = \frac{1}{2}\left(\frac{1}{6} + \frac{5}{6}\right) \times 2$

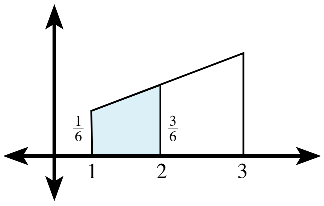
$= \frac{1}{2} \times \frac{6}{6} \times 2 = 1$



b $p(X \leqslant 2) = p(1 \leqslant X \leqslant 2)$

$= \frac{1}{2}\left(\frac{1}{6} + \frac{3}{6}\right) \times 1$

$= \frac{1}{2} \times \frac{4}{6} = \frac{4}{12} = \frac{1}{3}$



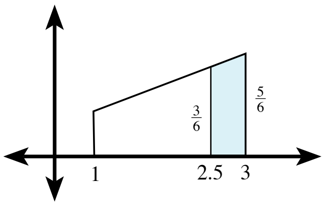
, $p(x > 2.5) = p(2.5 < x \leqslant 3)$

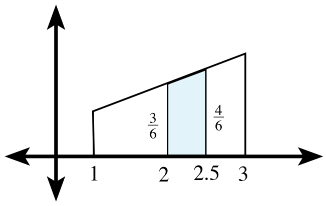
$$= \frac{1}{2}\left(\frac{4}{6} + \frac{5}{6}\right) \times \frac{1}{2}$$

$$= \frac{1}{2} \times \frac{9}{6} \times \frac{1}{2} = \frac{9}{24} = \frac{3}{8}$$

, $p(2 \leqslant x \leqslant 2.5) = \frac{1}{2}\left(\frac{3}{6} + \frac{4}{6}\right) \times \frac{1}{2}$

$$= \frac{1}{2} \times \frac{7}{6} \times \frac{1}{2} = \frac{7}{24}$$





Notice that : $p(2 \leqslant x \leqslant 2.5) = 1 - [\, p(x \leqslant 2) + p(x \geqslant 2.5)\,]$

$$= 1 - \left(\frac{1}{3} + \frac{3}{8}\right) = 1 - \frac{17}{24} = \frac{7}{24}$$

Try to solve

3 If x is a continuous random variable where :

$$f(x) = \begin{cases} \frac{1}{50}(17 - 2x) & \text{where } 1 < x < 6 \\ \text{zero} & \text{otherwise} \end{cases}$$

a **Prove that** f(x) is a probability density function of the random variable X .

b **Find** $p(x > 3)$

c **Find** $p(4 < X < 7)$

Example

2 If x is a continuous random variable whose probability density function is:

$$f(x) = \begin{cases} \frac{2x + k}{24} & \text{where } 1 < x < 4 \\ \text{zero} & \text{otherwise} \end{cases}$$

a Find the value of k.

b find $p(X > 3)$

Solution

$\because\ p(1 < x < 4) = 1$

$\therefore\ \frac{1}{2} \times 3 \times \frac{10 + 2k}{24} = 1$

$\therefore\ \frac{1}{2}\left(\frac{2 + k}{24} + \frac{8 + k}{24}\right) \times 3 = 1$

$\therefore\ k = 3$

, $f(3) = \frac{6 + 3}{24} = \frac{9}{24}$

$f(4) = \frac{8 + 3}{24} = \frac{11}{24}$,

$$\therefore\ P(x > 3) = \frac{1}{6}\left(\frac{9}{24} + \frac{11}{24}\right) \times 1 = \frac{1}{6} \times \frac{20}{24} = \frac{5}{12}$$

Try to solve

4 If X is a continuous random variable whose probability density function is :

$$f(x) = \begin{cases} \frac{2x + 1}{28} & 1 < x < 5 \\ \text{zero} & \text{otherwise} \end{cases}$$

a Find the value of a if $p(x < a) = \frac{1}{7}$

b Find the value of b if $p(b < x < b + 2) = \frac{1}{2}$

Exercises (3 - 3)

First : choose the correct answer :

1 If the probability distribution of the continuous random variable X is :

$$f(x) = \begin{cases} \frac{1}{2} & \text{where } 2 < x < 4 \\ \text{zero} & \text{otherwise} \end{cases}$$ then p (X > 3) =

a $\frac{1}{4}$ b $\frac{1}{2}$ c $\frac{3}{4}$ d 1

2 If the probability distribution of the continuous random variable X is :

$$f(x) = \begin{cases} k\,x & \text{where } 2 < x < 4 \\ \text{zero} & \text{otherwise} \end{cases}$$ then k =

a $\frac{1}{6}$ b $\frac{1}{3}$ c $\frac{1}{2}$ d $\frac{3}{4}$

3 If the probability distribution of the continuous random variable x is:

$$f(x) = \begin{cases} \frac{1}{6} & \text{where } -3 < x < 3 \\ \text{zero} & \text{otherwise} \end{cases}$$ then p (X = 3) =

a zero b $\frac{1}{6}$ c $\frac{1}{3}$ d $\frac{1}{2}$

Second : Answer the following questions :

4 If x is a continuous random variable where

$$f(x) = \begin{cases} \frac{x+3}{18} & \text{where } -3 < X < 3 \\ \text{zero} & \text{otherwise} \end{cases}$$

Find : **first**: p (X < 0) **second** : p (-1 < X < 2)

5 If x is a continuous random variable whose probability density function is:

$$f(x) = \begin{cases} \frac{2x+1}{24} & \text{where } 2 < x < 5 \\ \text{zero} & \text{otherwise} \end{cases}$$

find : **first** : p (3 < X < 5) **second** : p (x > 4)

6 If X is a continuous random variable where :

$$f(x) = \begin{cases} \frac{2\,(x+1)}{27} & \text{where } 2 < x < 5 \\ \text{zero} & \text{otherwise} \end{cases}$$

first : prove that f(x) is a density function to the random variable X.

second : Find p (X > 3)

7. If X is a continuous random variable whose probability density function is:

$$f(x) = \begin{cases} \frac{2x+1}{18} & \text{where } 1 < x < 4 \\ \text{zero} & \text{otherwise} \end{cases}$$

find : **first** : $p(X > 3)$ **second** : $p(2 < X < 4)$

8. If X is a continuous random variable whose probability density function is:

$$f(x) = \begin{cases} ax & \text{where } 0 < x < 4 \\ \text{zero} & \text{otherwise} \end{cases}$$

find : **first** : the value of a **second** : $p(1 < X < 3)$

9. If X is a continuous random variable whose probability density function is:

$$f(x) = \begin{cases} \frac{1}{8}x + \mathbf{a} & \text{where } 0 < x < 4 \\ \text{zero} & , \text{otherwise} \end{cases}$$

find : **first** : the value of a **second** : $p(1 < X < 3)$

10. If X is a continuous random variable whose probability density function is:

$$f(x) = \begin{cases} \frac{ax}{2} & \text{where } 0 < x < 4 \\ \text{zero} & \text{otherwise} \end{cases}$$

find : **first** : the value of a **second** : $p(1 < X < 3)$

11. If X is a continuous random variable whose probability density function is:

$$f(x) = \begin{cases} \frac{x-1}{k} & \text{where } 1 < x < 5 \\ \text{zero} & \text{otherwise} \end{cases}$$

find : **first** : the value of k **second** : $p(2 < X < 3)$

Critical thinking :

12. If X is a continuous random variable whose probability density function is:

$$f(x) = \begin{cases} \frac{x}{6} & \text{where } 0 < x < 2 \\ \frac{1}{3} & \text{where } 2 < x < 4 \\ \text{zero} & \text{otherwise} \end{cases}$$

Calculate : **a** $p(1 < x < 2)$ **b** the value of a making $p(2 < X < \mathbf{a}) = 0.5$

13. If X is a continuous random variable whose probability density function is:

$$f(x) = \begin{cases} \frac{3x+1}{40} & \text{where } 1 < x < 5 \\ \text{zero} & \text{otherwise} \end{cases}$$

and if **a**, b ∈ [1, 5[find

a the value of a if $p(a < X < a+2) = \frac{7}{20}$ **b** the value of b if $p(X > b) = \frac{69}{80}$

Unit Summary

1 The random variable is:

Function whose domain is the elements of the sample space and its co-domain is the set of real numbers R.

2 The discrete random variable :

Its range is a finite set. (i.e. it can be counted) of the set of real numbers.

3 The continuous random variable :

its range is an interval of the real numbers (closed or opened) i.e. it is a non-counted set of real numbers.

4 Probability distribution function of the discrete random variable :

If x is a discrete random variable whose range is the set:{ x_1 , x_2 , x_3. ...... , x_r } then the function f defined as follows: $f(x_r) = P(X = x_r)$ for each r = 1 , 2 , 3 , ...

Identifies what is called the probability distribution function of the discrete random variable X which is expressed by a set of ordered pairs specifying the function F.

i.e. the probability distribution of the random variable X = { (x_1 , $f(x_1)$) , (x_2 , $f(x_2)$) , (x_3 , $f(x_3)$) , ... , (x_n , $f(x_n)$) }

5 Expectation (mean):

Expectation is the value at which the most values of the discrete random variable are centralized and sometimes is called " mean" and is denoted by the symbol (μ) and read as (Mu).

If X is a discrete random variable whose probability distribution function is f and its range is: { x_1 , x_2 , x_3 , ... , x_n} by probabilities of $f(x_1)$, $f(x_2)$, $f(x_3)$, ... , $f(x_n)$ respectively, then the expectation is given by the relation:

$$\text{Expectation } (\mu) = \sum_{r=1}^{n} x_r \times f(x_r)$$

i.e.: Expectation (μ) = $x_1 \times f(x_1) + x_2 \times f(x_2) + x_3 \times f(x_3) + ... + x_n \times f(x_n)$

6 variance:

The variance of a discrete random variable measures the amount of dispersion of the discrete random variable of its expected value and is denoted by the symbol (σ^2) and read as (squared sigma) and is given by the relation

$$\sigma^2 = \sum_{r=1}^{n} x_r^2 \times f(x_r) - \mu^2$$

Note: The standard deviation of a discrete random variable x is the square root of the variance and is denoted by the symbol σ . It's noticed that the variance and the standard deviation are always positive quantities .

7 Coefficient of Variation:

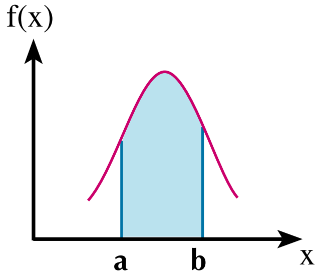
The coefficient of variation for any set of items is known as the percentage between the standard deviation of the set and its expectation (mean) and it can be determined as in the following relation :

$$\textbf{Coefficient of variation} = \frac{\textbf{standard deviation}}{\textbf{mean}} \times 100\,\% = \frac{\sigma}{\mu} \times 100\%$$

This coefficient forms the dispersion of the set in the form of a percentage abstracted from distinction where it is not affected by the measured units of the phenomenon.

8 Probability Density Function :

For any continuous random variable x, there is a real function whose range is non- negative, and denoted by the symbol f(x) called the probability density function by which the probabilities of events expressing it by the random variable can be found throughout the area included under the function curve and above the x axis and p (a<x<b) can be calculated by finding the area of the shaded part of the curve of the function f between the two values a and b as shown in the opposite figure .



This function satisfies the next conditions:

- f(x) ⩾ 0 for all the values of x belonging to the domain of the function.
- The area of the region located under the curve of the function f and above the x axis equal 1.

General exercises

First : Choose the correct answer :

1. The expected value of the following probability distribution is:

x_r	zero	1	2
$f(x_r)$	0.2	0.3	0.5

a) 1 b) 1.14 c) 1.3 d) 1.5

2. If the expectation of the following probability distribution is:

x_r	1	2	k
$f(x_r)$	0.1	0.8	0.1

equals 2, then the value of k equals

a) 3 b) 4 c) 5 d) 6

3. If x is a continuous random variable whose probability density function is :

$$f(x) = \begin{cases} k & \text{where } -4 < x < 4 \\ \text{zero} & \text{otherwise} \end{cases} \quad \text{then } k =$$

a) $\frac{1}{8}$ b) $\frac{1}{4}$ c) zero d) 4

second: answer the following questions

4. If X is a discrete random variable whose range $= \{-3, -1, 0, 1, 3\}$ and the values of $p(X = -3) = p(X = 3) = \frac{1}{9}$, $p(X = 0) = \frac{4}{9}$, $p(X = -1) = p(X = 1)$ find the probability distribution function of the discrete random variable.

5. If X is a discrete random variable whose probability distribution function is determined by the function $f(x) = \frac{x+4}{16}$ where $x = -2, m, 1, 2$ find the value of the constant m, then write down he probability distribution function of the variable X

6. If X is a discrete random variable whose probability distribution is determined by the function $f(x) = \frac{a x^2}{2}$ where $x = 0, 1, 2, 3$ find the value of a then find $p(X \leqslant 2)$.

7. In an experiment of rolling a regular die two consecutive times and observing the number on the upper face in each time. Find the probability distribution of the discrete random variable expressing the greater number appeared on the upper two faces.

8. Design a die in a way two faces carry the number 1, two faces carry the number 3 and two faces carry the number 5. Roll this die two consecutive times and observe the number appearing the upper face in each time. If the discrete random variable X expresses "the absolute difference between the two upper numbers"

Find : First : the probability distribution of the discrete random variable X.

Second : the probability of getting the absolute difference between the two numbers less than 4 .

9. Two regular dices; each two opposite faces of the first die are numbered {1, 3, 5} and each two opposite faces of the second die are numbered { 2 , 4 , 6 } If the two dices have been rolled and the random variable X expresses the sum of the two numbers, find the probability distribution function of the variable x and calculate the mean and the standard deviation of the variable x.

10. Two boxes a and b, each contains 4 balls numbered from 1 to 4. A ball is randomly drawn from each box. If the random variable is " the sum of the two numbers appeared on the two drawn balls", find the probability distribution function of the variable X and calculate the mean and the standard deviation.

11. If x is a random variable whose mean $\mu = 2$ and its probability distribution is as follows:

x_r	1	0	2	a
$f(x_r)$	$\frac{1}{12}$	b	$\frac{1}{3}$	$\frac{5}{12}$

First : calculate the two values of a and b **Second** : calculate the standard variation

12. If X is a continuous random variable whose probability density function:

$$f(x) = \begin{cases} \frac{x+2}{16} & \text{Where } 0 < x < 4 \\ \text{Zero} & \text{otherwise} \end{cases}$$

Find : First : P (X < 2) **Second** : p (1 < X < 4)

13. If X is a continuous random variable whose probability density function :

$$f(x) = \begin{cases} \frac{x+1}{8} & \text{Where } a < x < a+2 \\ \text{Zero} & \text{otherwise} \end{cases}$$

Find : First : the value of a **Second** : p (X < 3)

For more activities and exercises, visit www.sec3mathematics.com.eg

Accumulative test

1. In the spinning wheel game, the wheel has been divided into 16 identical sectors numbered from 1 to 16. What is the probability the indicator lands on an odd number known that it is landed on a number greater than 5?

2. **Roads** : the next table shows the probability distribution for the number of the accidents expected during the rain days on the roads.

Number of accidents	zero	1	2	3	4	5	6
Probability	0.1	0.26	0.31	0.14	0.11	0.06	0.02

Calculate the expected value for the number of these accidents.

3. **Media** : Some electronic sites have formed a survey related to the most TV program seen by people and the results have been recorded as shown in the following table :

Type of the program	Cultural	Social	News	Sports	Entertaining	Others
Probability of occurrence	0.14	0.2	0.24	0.18	0.16	0.08

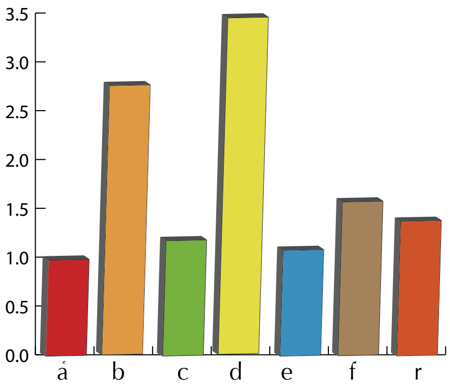
a Represent these data by columns

b Prove that these data represent a probability distribution

c If a spectator has been randomly chosen, find the probability of choosing a social or sports program.

d Do a report about the effects media on forming the social culture.

4. **Sports** : 7 racers have participated in a short-range race, the probability to win such a race has been represented in the opposite graph :



a Show which of these distributions represents a probability distribution.

b Find the probability racers b, a, and e to win this competition.

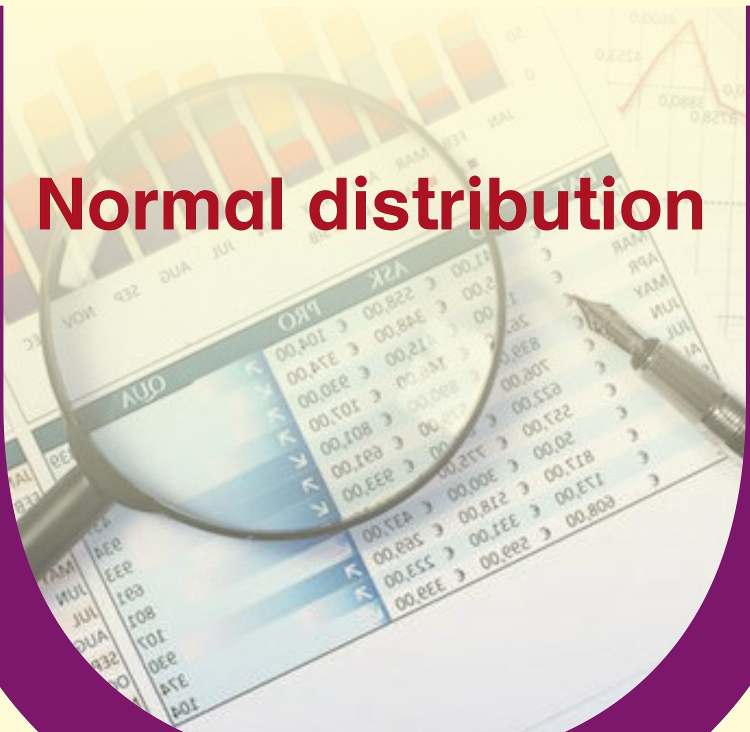
5. If X is a continuous random variable whose probability density function is :

$$f(x) = \begin{cases} \dfrac{x+1}{12} & \text{where } 0 < x < 4 \\ \text{zero} & \text{otherwise} \end{cases}$$

Find : **first** : $p(X < 2)$ **Second** : $p(2 < X < 5)$

Unit 4

Normal distribution



introduction

The normal distribution is considered one of the most important probability distributions learned in the statistics curricula due to its various uses to the outcomes of some processes in the physical, economic and social sciences and it deals with the phenomena of our daily life. The french scientist Abraham de Moivre was the first to use the normal distribution in 1756 in one of its prints. Some other scientists such as the German scientist Carl Friedrich Gauss(1777-1855) has participated in developing the normal distribution. The normal distribution is sometimes named after him (Gauss curve).

Carl Friedrich Gauss

Abraham de Moivre

The most well-known application of the normal distribution is the administrative evaluation of the subordinates in order to ensure the justice. The normal distribution is used to study the remains of analyzing the slope and it has a close relation with the control charts and so on.

Unit objectives

By the end of the unit and carrying out the involved activities, the student should be able to:

- Identify the normal distribution and its properties.
- Calculate the probability of the standard variable.
- Calculate the probability of the non-standard normal variable.
- Identify the standard normal random variable and the general form of the curve representing the density function of this variable.
- Convert any normal random variable into a standard normal variable.
- Find the values of the probabilities of a random variable with a standard normal distribution using the statistical tables.
- Explain the properties of the normal distribution curve and some phenomena it expresses.
- Explain the results which they have obtained through calculating the probability of the normal random variable.

Key Terms

- Normal distribution
- Normal random variable
- Normal curve
- Standard normal curve

Lessons of the unit

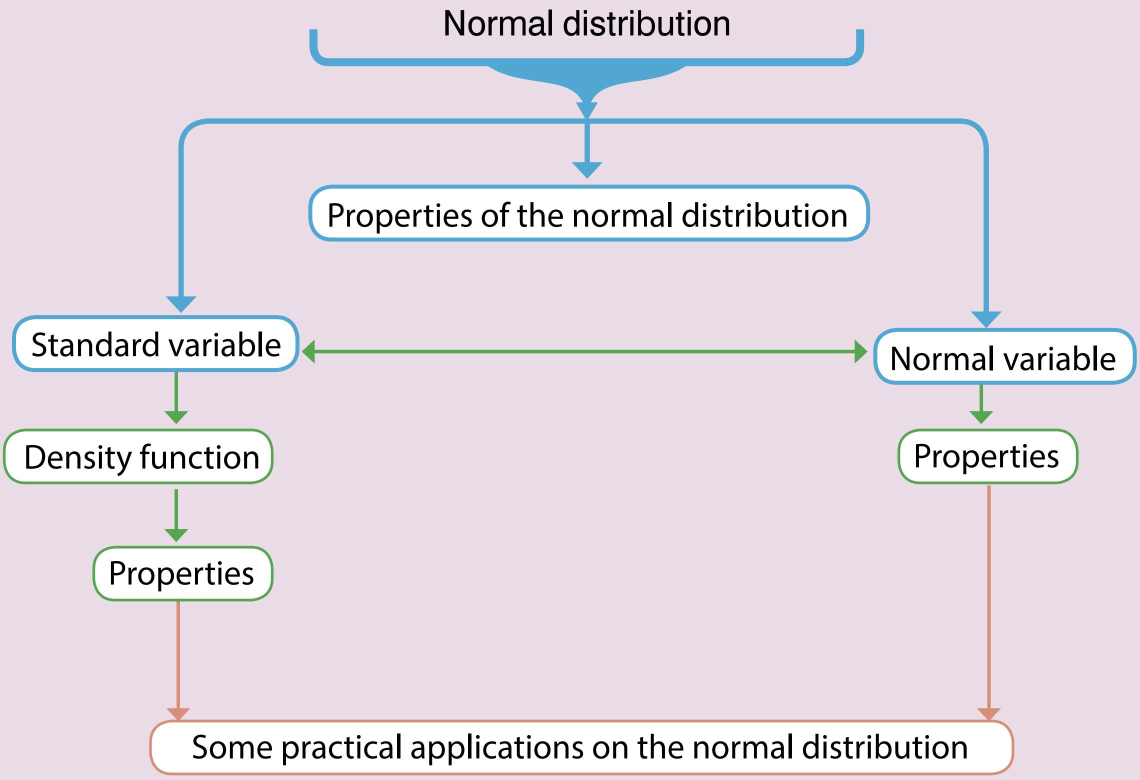
Lesson (4-1) Normal distribution

Lesson (4-2) Some practical applications on the normal distribution

Materials

Scientific calculator

Unit planning organizer



Unit 4

4 - 1 Normal distribution

You will learn

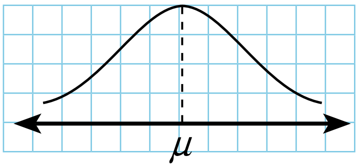
- Normal random variable
- Some properties of the normal distribution
- Standard normal distribution
- Properties of the density function of the standard normal distribution
- Calculating the probability of the standard normal distribution.

Key terms

- Normal distribution
- Normal random variable
- Normal curve
- Standard normal distribution

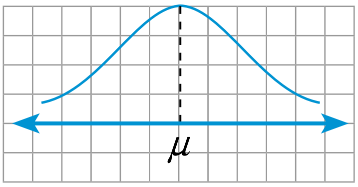
Introduction:

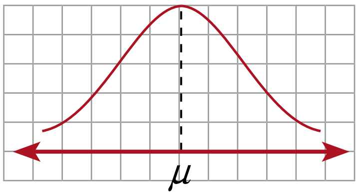
The normal distribution is one of the most important continuous probability distributions since it has important theoretical properties. Furthermore, the outcomes of the normal distribution can take any value within an interval of the real numbers. For example, the lengths of adults, the weights of babies as they were born and the intelligence degree of people and so on. The normal distribution can be described as a mathematical equation identifying its curve. This equation is identified thoroughly by knowing the expectation (mean) μ and the standard deviation σ . This curve is bell-like. It is symmetrical on the straight line x = μ and its two ends are converged from the horizontal axis where its two ends expand infinitely as shown in the opposite figure.



Normal random variable:

It is said that the continuous random variable x is "a normal random variable" if its range can be identified by the interval]-∞ , ∞[and its probability density function is represented by a bill-like curve. The curve of the density function is called the normal curve or "Gauss curve". The value of the curve is identified by knowing two basic values: the mean μ and the standard deviation σ of the random variable x as shown in the next figures.





Some properties of the normal curve

(1) It has one top and its two ends extends to - ∞ , ∞ .

(2) It has a symmetrical axis passing through the top and intersecting the horizontal axis when X = μ .

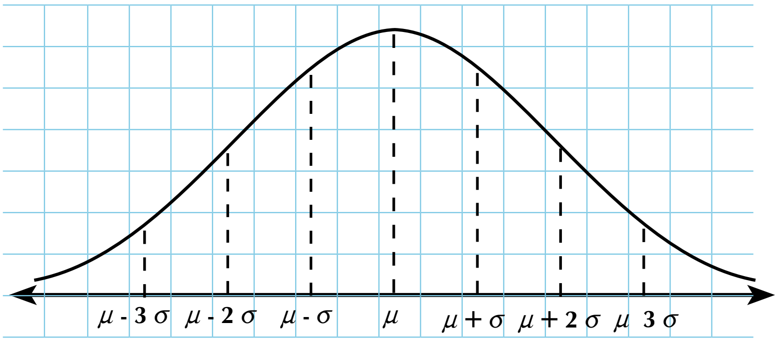
(3) The area of the region placed under the normal curve and above the X axis equals 1 .

(4) From symmetry, we find that the straight line X = μ divides the area placed under the curve and above the X axis into two regions; the area of each region = 0.5.

Materials — Scientific calculator.

(5) The approximate area of the region under the curve and above the X axis can be calculated in regard to the following intervals:

- From $\mu - \sigma$ to $\mu + \sigma = 68.26\ \%$ the total area .
- From $\mu - 2\sigma$ to $\mu + 2\sigma = 95.44\ \%$ the total area .
- From $\mu - 3\sigma$ to $\mu + 3\sigma = 99.74\ \%$ the total area .



Notice : The number of data should be big in order to make the normal distribution approximated

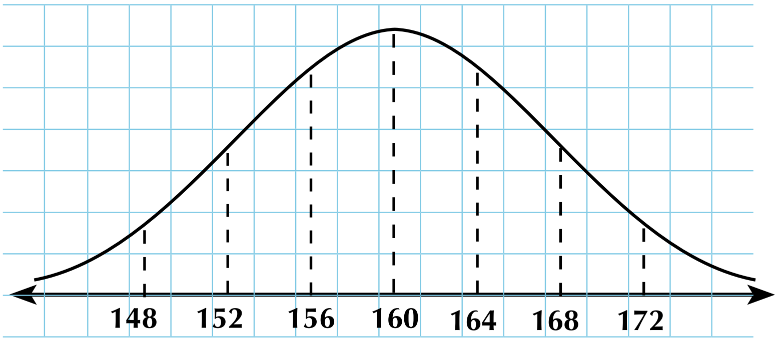
Example

1 If the lengths of students at a school follow a normal distribution of an average of 160 cm, the standard deviation is 4 cm. a student is randomly chosen, find the probability of the length of the student is:

a Taller than 172 cm   **b** Shorter than 156 cm

c Included between 156 cm and 168 cm

Solution



From the given data, we find that: the mean $\mu = 160$ and the standard deviation $\sigma = 4$

By comparing the data with the normal distribution curve, we find that: $\mu + 3\text{ x} = 160 + 3 \times 4$ so,

a $P(x > 172) = P(x > \mu + 3\sigma)$

$\because$ The area from $\mu - 3\sigma$ to $\mu + 3\sigma = 0.9974$

$\therefore$ The area from μ to $\mu + 3\sigma = 0.9974 \div 2 = 0.4987$

$\therefore$ The area on the right of $\mu + 3\sigma = 0.5 - 0.4987 = 0.0013$

b $P(x < 156) = P(x < \mu - \sigma)$

$\because$ The area from $\mu - \sigma$ to $\mu + \sigma = 0.6826$

$\therefore$ The area from μ to $\mu - \sigma = 0.6826 \div 2 = 0.3413$

$\therefore$ The area on the left of $\mu - \sigma = 0.5 - 0.3413 = 0.1587$

c $P(156 < x < 168) = P(\mu - \sigma < x < \mu + 2\sigma) = P(\mu - \sigma < x < \mu) + P(\mu < x < \mu + 2\sigma)$

$$= \frac{0.6816}{2} + \frac{0.9544}{2} = 0.3408 + 0.4772 = 0.818$$

Try to solve

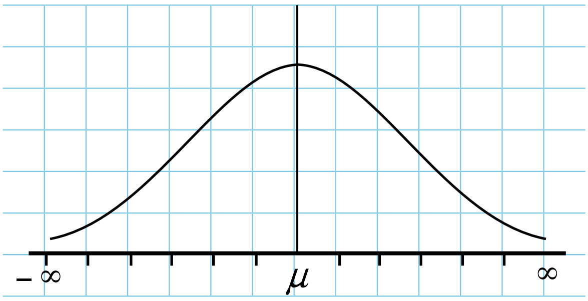
1 If the weights of the students in a faculty follow a normal distribution μ of an average = 68 kg and its variance is 16 kg^2, find:

a The probability that the weight is greater than 72 kg

b The percentage of the students whose weights range between 64 kg and 72 kg "the weight of each student".

c The number of the students whose weights are more than 64 kg if the total number of the college students is 2000 students.

The standard normal distribution

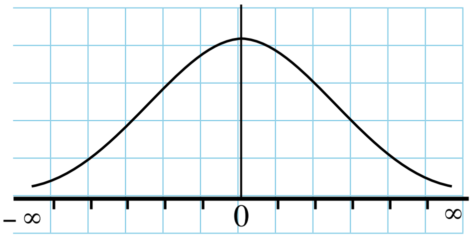


In the normal distribution, we have noticed that when you find the probability, the lengths of the intervals are multiples of the standard deviation in order to calculate the probability. Hence, it has been proper to convert the normal distributions into standard normal distributions by converting the values of (x) into standard values of (z) in terms of the mean μ and the standard deviation σ; at which $\mu = 0$, $\sigma = 1$

Definition

If the probability distribution of the random variable x is the normal distribution by a mean μ and standard deviation σ, then $z = \frac{x - \mu}{\sigma}$ is the standard normal distribution whose mean is μ = zero and standard deviation $\sigma = 1$.

Some properties of the density function of the standard normal distribution (z).



(1) The curve is located above the horizontal axis (X axis).

(2) It is symmetrical in terms of the vertical axis (Y axis).

(3) The two ends of the curve extends infinitely without getting converged with the horizontal axis.

(4) The area of the region under the curve and above the horizontal axis = 1.

(5) From the symmetry, we find that the vertical axis divides the region under the curve and above the horizontal axis into two regions the area of each = 0.5.

(6) The approximate area of the region located under the standard curve only and above any interval]a , b[by special tables can be calculated.

The table of the area under the standard normal distribution curve:

To convert the normal distribution X into standard normal distribution Z you use the relation:

$Z = \dfrac{x - \mu}{\sigma}$, from the table of the standard normal distribution attached at the end of the book, you can find the required area .

Here, we explain how you can use the table of area under the standard normal distribution.

Z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0						0.0199				
0.1										
0.2										
0.3										
0.4	0.1554									
0.5										
0.6				0.2357						
2.5								0.4949		
3.5										

$P(0 \leqslant Z \leqslant 0.05)$ = the area under the standard normal distribution is above the interval [0 , 0.05] i.e. z = 0.05, and so, we search in the table in raw 0.00 and under the column 0.05, to find the number is 0.0199

$\therefore P(0 \leqslant Z \leqslant 0.05) = 0.0199$

$P(0 \leqslant Z \leqslant 0.4)$ = the area under the standard normal distribution is above the interval [0 , 0.4] i.e. z = 0.4 , and so, we search in the table in raw 0.4 and under the column 0.00, to find the number is 0.1554.

$\therefore P(0 \leqslant Z \leqslant 0.4) = 0.1554$

$P(0 \leqslant Z \leqslant 0.63)$ = the area under the standard normal distribution is above the interval [0 , 0.63] I.e. z = 0.63 , and so, we search in the table in raw 0.6 and under the column 0.03, to find the number is 0.2357. 0.2357

$\therefore P(0 \leqslant Z \leqslant 0.63) = 0.2357$

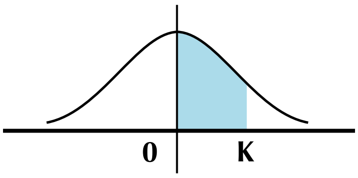
$P(0 \leqslant Z \leqslant 2.57)$ = the area under the standard normal distribution is above the interval [0 , 2.57] I.e. z = 2.57 , and so, we search in the table in raw 2.5 and under the column 0.07, to find the number is 0.4949

$\therefore P(0 \leqslant Z \leqslant 2.57) = 0.4949$

Calculating the probability of the standard normal variable:

(1) To find the area of the region under the curve in the interval [0 , K] from the table

The table of areas under the standard normal curve gives the approximate area above the interval [0 , K] and under the normal curve where $K \geqslant 0$, i.e. the table directly gives us $P(0 \leqslant Z \leqslant K)$



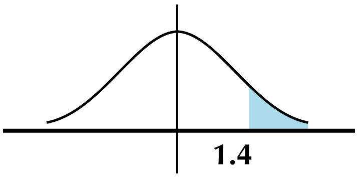
For example: $P(0 \leqslant Z \leqslant 0.3) = 0.1179$, $P(0 < Z \leqslant 0.64) = 0.2389$

$P(0 \leqslant Z \leqslant 1.7) = 0.4554$, $P(0 \leqslant Z < 2.45) = 0.4929$

Notice that: $P(Z \geqslant 1.4) = 0.5 - P(0 \leqslant Z \leqslant 1.4)$

$= 0.05 - 0.4192$

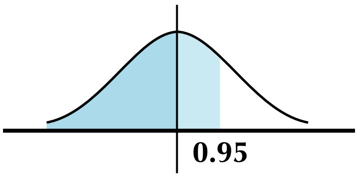
$= 0.0808$



Similarly: $P(Z \leqslant 0.95) = 0.5 + P(0 \leqslant Z \leqslant 0.95)$

$= 0.5 + 0.3289$

$= 0.8289$



(2) Finding the area of the region under the curve in the interval [- K , 0] from the table

From the symmetry of the standard normal curve around the vertical axis, we find that :

$P(-K \leqslant Z \leqslant 0) = P(0 \leqslant Z \leqslant K)$

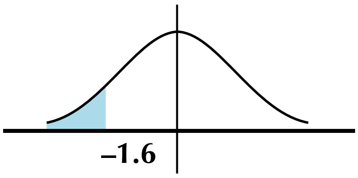
For example: $P(-1.25 \leqslant Z \leqslant 0) = P(0 \leqslant Z \leqslant 1.25) = 0.3944$

, $P(-2.24 \leqslant Z \leqslant 0) = P(0 \leqslant Z \leqslant 2.24) = 0.4875$

, $P(Z \leqslant -1.6) = 0.5 - P(-1.6 \leqslant Z \leqslant 0)$

$= 0.5 - P(0 \leqslant Z \leqslant 1.6)$

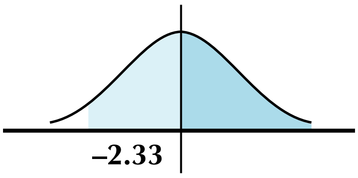
$= 0.5 - 0.4452 = 0.0548$



, $P(Z \geqslant -2.32) = 0.5 + P(-2.32 \leqslant Z \leqslant 0)$

$= 0.5 + P(0 \leqslant Z \leqslant 2.32)$

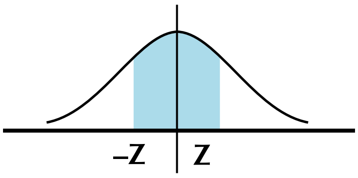
$= 0.5 + 0.4898 = 0.9898$



Notice that: $P(-K \leqslant Z \leqslant K) = 2 \times P(0 \leqslant Z \leqslant K)$

For example: $P(-1.4 \leqslant Z \leqslant 1.4) = 2 \times P(0 \leqslant Z \leqslant 1.4)$

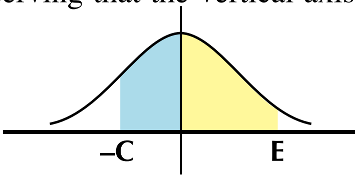
$= 2 \times 0.4192 = 0.8384$



, $P(-2.0 \leqslant Z \leqslant 2.0) = 2 \times P(0 \leqslant Z \leqslant 2.0) = 2 \times 0.4772 = 0.9544$

(3) Finding the area of the region under the curve in any interval [C , E]:

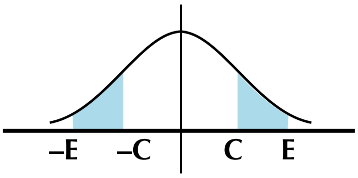
In this case, it is favorable to draw the standard curve and observing that the vertical axis divides the area under the curve and above the horizontal axis into two equal areas and the area of each = 0.5.



First: P (- C ⩽ Z ⩽ E) where C and E are positive.

= P (- C ⩽ Z ⩽ 0) + P (0 ⩽ Z ⩽ E)

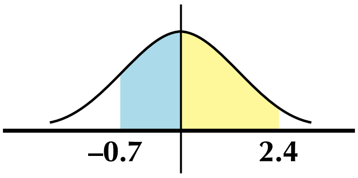
= P (0 ⩽ Z ⩽ c) + P (0 ⩽ Z ⩽ E)



Second: P (C ⩽ Z ⩽ E) = P (- E ⩽ Z ⩽ -C)

= P (0 ⩽ Z ⩽ E) – P (0 ⩽ Z ⩽ C)

For example:

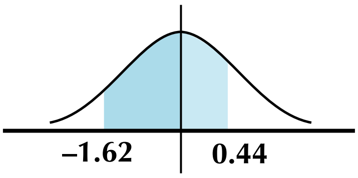


(1) P (-0.7 ⩽ Z ⩽ 2.4)

= P (-0.7 ⩽ Z ⩽ 0) + P(0 ⩽ Z ⩽ 2.4)

= P (0 ⩽ Z ⩽ 0.7) + P(0 ⩽ Z ⩽ 2.4) from symmetry

= 0.2580 + 0.4918 = 0.7498

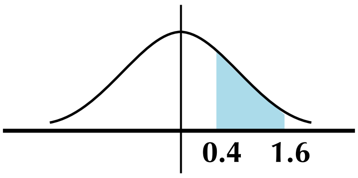


(2) P (-1.62 < Z ⩽ 0.44)

= P(-1.62 < Z ⩽ 0) + P(0 ⩽ Z ⩽ 0.44)

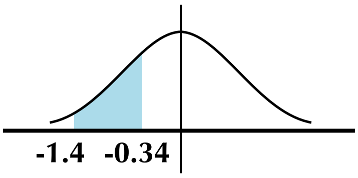
= P (0 ⩽ Z ⩽ 1.62) + P(0 ⩽ Z ⩽ 0.44) from symmetry

= 0.4474 + 0.1700 = 0.6174



(3) P (0.4 ⩽ Z < 1.6) = P(0 ⩽ Z < 1.6) - P (0 ⩽ Z ⩽ 0.4)

= 0.4452 – 0.1554 = 0.2898



(4) P (- 1.4 < Z < -0.34)

= P(- 1.4 < Z ⩽ 0) - P(-0.34 < Z ⩽ 0)

= P (0 ⩽ Z ⩽ 1.4) + P(0 ⩽ Z ⩽ 0.34) from symmetry

= 0.4192 – 0.1331 = 0.2861

Example Finding the area under the standard normal curve

2 If Z is a standard normal random variable, find:

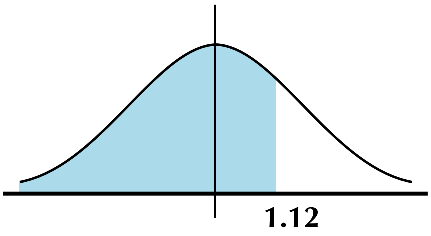
a P(Z ⩽ 1.12) **b** P(Z ⩾ 1.64) **c** P(0.48 ⩽ Z ⩽ 2.1)

Solution

a P(Z ⩽ 1.12)

= P(0 ⩽ Z ⩽ 1.12) + P(Z ⩽ 0)

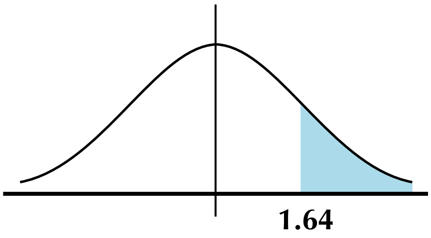
= 0.3686 + 0.5 = 0.8686



b P(Z ⩾ 1.64)

$= P(Z \geqslant 0) - P(0 \leqslant Z \leqslant 1.64)$

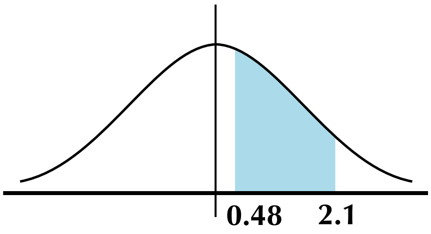
$= 0.5 - 0.4495 = 0.0505$



c P(0.48 ⩽ Z ⩽ 2.1)

$= P(0 \leqslant Z \leqslant 2.1) - P(0 \leqslant Z \leqslant 0.48)$

$= 0.4821 - 0.1844 = 0.2977$



Try to solve

2 If Z is a standard normal random variable, find:

a $P(0 \leqslant Z \leqslant 0.82)$ **b** $P(Z \geqslant 2.32)$

c $P(Z \leqslant 1.64)$ **d** $P(1.08 \leqslant Z \leqslant 3.12)$

Example

3 If Z is a standard normal random variable, find:

a $P(Z \leqslant -0.56)$ **b** $P(Z \geqslant -1.06)$

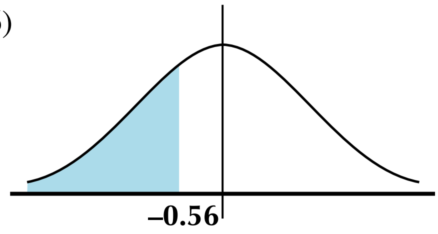
c $P(-1.2 \leqslant Z \leqslant 2.48)$ **d** $P(-2.2 \leqslant Z \leqslant -0.46)$

Solution

a $P(Z \leqslant -0.56)$

$= P(Z \geqslant 0.56)$

$= 0.5 - P(0 \leqslant Z \leqslant 0.56) = 0.5 - 0.2123 = 0.2877$

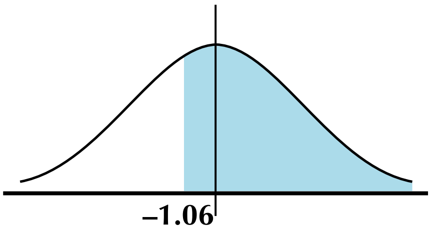


b $P(Z \geqslant -1.06)$

$= P(Z \leqslant 1.06)$

$= P(0 \leqslant Z \leqslant 1.06) + 0.5$

$= 0.0636 + 0.5 = 0.5636$

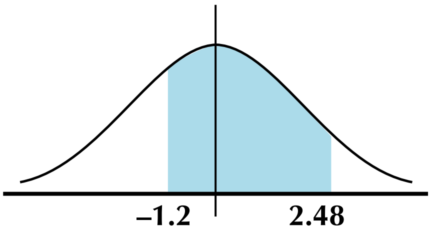


c $P(-1.2 \leqslant Z \leqslant 2.48)$

$= P(-1.2 \leqslant Z \leqslant 0) + P(0 \leqslant Z \leqslant 2.48)$

$= P(0 \leqslant Z \leqslant 1.2) + P(0 \leqslant Z \leqslant 2.48)$

$= 0.3849 + 0.4934 = 0.8783$

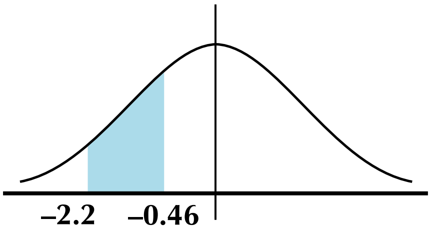


d $P(-2.2 \leqslant Z \leqslant -0.46)$

$= P(-2.2 \leqslant Z \leqslant 0) - P(-0.46 \leqslant Z \leqslant 0)$

$= P(0 \leqslant Z \leqslant 2.2) - P(0 \leqslant Z \leqslant 0.46)$

$= 0.4861 - 0.1772 = 0.3089$



Try to solve

3 If Z is a standard normal random variable, find:

a $P(Z \leqslant -0.56)$ **b** $P(Z \geqslant -1.06)$

c $P(-1.2 \leqslant Z \leqslant 2.48)$ **d** $P(-2.2 \leqslant Z \leqslant -0.46)$

Example Converting from a normal variable into a standard normal variable

4 If X is a normal random variable whose mean is μ and standard deviation is σ , find:

a $P(X > \mu - 1.5\sigma)$ **b** $P(X < \mu - 0.5\sigma)$

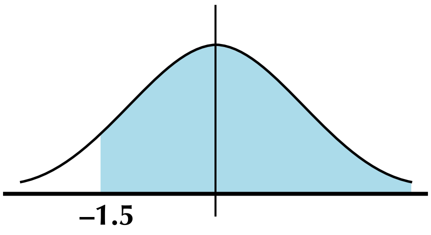
c $P(\mu - 1.96\sigma < X < \mu + 1.96\sigma)$

Solution

a $P(Z > \frac{\mu - \sigma 1.5 - \mu}{\sigma}) = P(Z > -1.5)$

$= P(-1.5 < Z < 0) + 0.5$

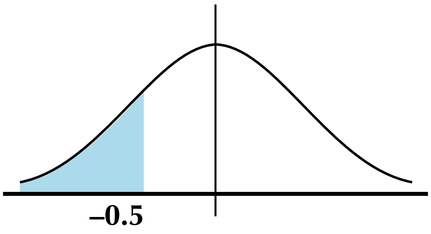
$= P(0 < Z < 1.5) + 0.5 = 0.4332 + 0.5 = 0.9332$



b $Z(X < \mu - 0.5\sigma) = P(Z < \frac{\mu - \sigma 0.5 - \mu}{\sigma})$

$= P(Z < -0.5) = P(Z > 0.5)$

$= 0.5 - P(0 < Z < 0.5) = 0.5 - 0.1915 = 0.3085$

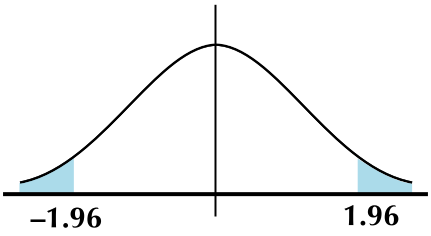


c $P(\mu - 1.96\sigma < X < \mu + 1.96\sigma)$

$= P(\frac{\mu - \sigma 1.96 - \mu}{\sigma} < Z < \frac{\mu - \sigma 1.96 - \mu}{\sigma})$

$= P(-1.96 < Z < 1.96)$

$= 2\,P(0 < Z < 1.96) = 2 \times 0.4750 = 0.95$



Try to solve

4 If X is a normal random variable whose mean is μ and standard deviation is σ, find:

a $P(X < \mu - 2.1\sigma)$ **b** $(X > \mu + 0.8\sigma)$

c $P(\mu - 1.48\sigma < X < \mu + 1.48\sigma)$

Example

5 If Z is a standard normal random variable, find the value of k in each of the following cases :

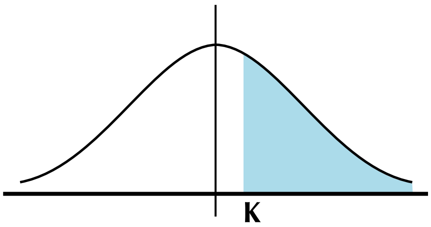
a $P(Z \geqslant K) = 0.1056$ **b** $P(Z \leqslant K) = 0.1151$

c $P(-0.44 \leqslant Z \leqslant K) = 0.5588$ **d** $P(K \leqslant Z \leqslant 2.1) = 0.2906$

Solution

a We notice that:

The area < 0.5 and the sign of the inequality "greater than", so K lies at the positive interval as shown in the opposite figure.



$\because P(Z \geqslant K) = 0.1056$

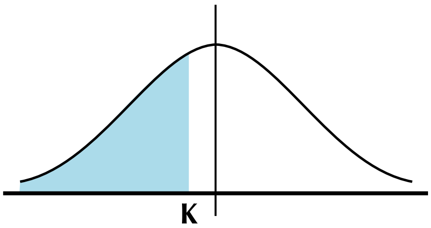
$\therefore 0.5 - P(0 \leqslant Z \leqslant K) = 0.1056$

$\therefore P(0 \leqslant Z \leqslant K) = 0.5 - 0.1056 = 0.3944$

In the area tables, we try to find the number (Z) or the nearest number corresponding the area 0.3944. We find that 1.2 is under the differences 0.05. **I.e.** **K = 1.25**

b We notice that:

The area < 0.5 , and the sign of the inequality "less than", so k lies at the negative interval as shown in the opposite figure.



$\because P(Z \leqslant K) = 0.1151$

From the symmetry in the curve, we find that: $P(Z \geqslant K) = 0.1151$

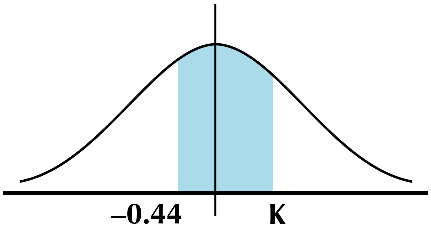
$\therefore 0.5 - P(0 \leqslant Z \leqslant K) = 0.1151$

$\therefore P(0 \leqslant Z \leqslant K) = 0.5 - 0.1151 = 0.3849$

$\therefore$ **K = - 1.2** (**notice that k lies at the negative part**)

c We notice that :

The area > 0.5 and an end of the interval lie at the negative interval, so the other end of the interval z lies at the positive interval as shown in the opposite figure.



$\because P(-0.44 \leqslant Z \leqslant K) = 0.5588$

$\therefore P(-0.44 \leqslant Z \leqslant 0) + P(0 \leqslant Z \leqslant K) = 0.5588$

$\therefore P(0 \leqslant Z \leqslant 0.44) + P(0 \leqslant Z \leqslant K) = 0.5588$

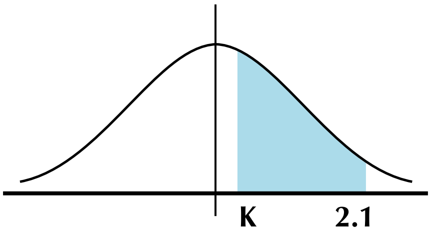
$\therefore 0.1700 + P(0 \leqslant Z \leqslant K) = 0.5588$

$\therefore P(0 \leqslant Z \leqslant K) = 0.5588 - 0.1700 = 0.3888$

$\therefore$ **K = 1.22**

d We notice that :

The area < 0.5 and an end of the interval lie at the positive interval, so the other end of the interval g lies at the positive interval as shown in the opposite figure. (also $P(0 \leqslant Z \leqslant 2.1) > 0.2906$)



$\because P(K \leqslant Z \leqslant 2.1) = 0.2906$

$\therefore P(0 \leqslant Z \leqslant 2.1) - P(0 \leqslant Z \leqslant K) = 0.2906$

$\therefore P(0 \leqslant Z \leqslant K) = P(0 \leqslant Z \leqslant 2.1) - 0.2906$

$= 0.4821 - 0.2906 = 0.1915$ $\therefore$ **K = 0.5**

Try to solve

5 If Z is a standard normal random variable, find the value of k in each of the following cases:

a $P(Z \geqslant K) = 0.1980$

b $P(Z \leqslant K) = 0.1980$

c $P(-2.4 \leqslant Z \leqslant K) = 0.7970$

d $P(K \leqslant Z \leqslant 2.5) = 0.8238$

Example

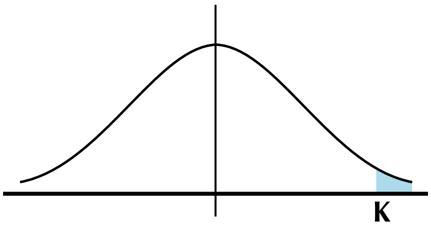
6 If x is a normal random variable whose mean is μ and standard deviation is σ

a	If: $P(X \geqslant 180)$	$= 0.0062$	, $\mu = 165$	calculate σ
b	If: $P(X > 35)$	$= 0.8643$	, $\sigma = 5$	calculate μ
c	If: $P(X \leqslant 170)$	$= 0.0228$	, $\sigma = 7$	calculate μ
d	If: $P(X \leqslant K)$	$= 0.8944$	, $\mu = 125$, $\sigma = 8$	calculate K
e	If: $P(X > K)$	$= 0.9452$	, $\mu = 50$, $\sigma = 5$	calculate K

Solution

a $P(X \geqslant 180) = P(Z \geqslant \frac{180 - 165}{\sigma}) = 0.0062$

$\therefore P(Z \geqslant K) = 0.0062$ where $K = \frac{15}{\sigma}$, $K > 0$

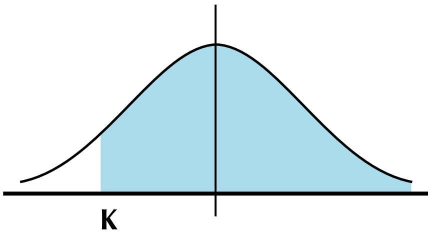


$\therefore P(0 \leqslant Z \leqslant K) = 0.5 - 0.0062 = 0.4938$

$\therefore K = 2.5$

$\therefore \frac{15}{\sigma} = \frac{5}{2}$ $\therefore \sigma = \frac{2 \times 15}{5}$ $\therefore$ **$\sigma = 6$**

b $P(X > 35) = P(Z > \frac{35 - \mu}{5}) = 0.8643$



$\therefore\ P(Z > K) = 0.8643$ where

$\therefore\ K = \frac{35 - \mu}{5}$, $K < 0$

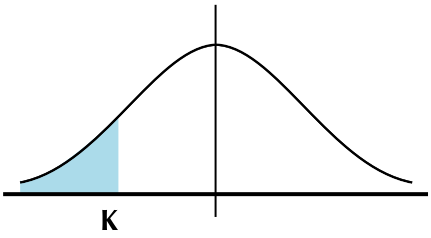
$\therefore\ P(0 \leqslant Z \leqslant -K) + 0.5 = -0.8643$

$P(K < Z \leqslant 0) = 0.8643 - 0.5 = 0.3643$ $\quad\therefore K = -1.1$

$\therefore \frac{35 - \mu}{5} = -1.1$ $\quad\therefore 35 - \mu = -5.5$

$\therefore \mu = 35 + 5.5$ $\quad\therefore$ **$\mu = 40.5$**

c $P(X \leqslant 170) = P(Z \leqslant \frac{170 - \mu}{7}) = 0.0228$



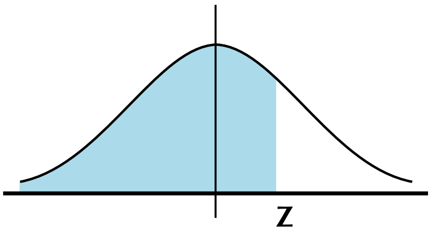
$\therefore P(Z \leqslant K) = 0.0228$ where $K = \frac{170 - \mu}{7}$, $K < 0$

$\therefore P(K \leqslant Z \leqslant 0) = 0.5 - 0.0228 = 0.4772$

$\therefore K = -2$

$\therefore \frac{170 - \mu}{7} = -2$ $\quad\therefore 170 - \mu = -14$ $\quad\therefore \mu = 170 + 14$ $\quad$ **$\mu = 184$**

d $P(X \leqslant K) = P(Z \leqslant \frac{K - 125}{8}) = 0.8944$



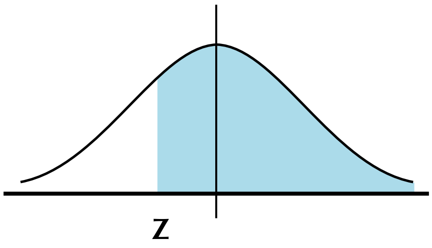
$\therefore P(Z \leqslant z) = 0.8944$

where $z = \frac{K - 125}{8}$, $z > 0$

$\therefore P(0 \leqslant Z \leqslant z) = 0.8944 - 0.5 = 0.3944$ $\quad\therefore z = 1.25$

$\therefore \frac{K - 125}{8} = 1.25$ $\quad\therefore K - 125 = 10$ $\quad\therefore K = 125 + 10$ $\quad\therefore K = 135$

e $P(X > K) = P(Z > \frac{K - 50}{5}) = 0.9452$



$\therefore P(Z > z) = 0.9452$

where $z = \frac{K - 50}{5}$, $z < 0$

$\therefore P(0 \leqslant Z < -z) = 0.9452 - 0.5 = 0.4452$ $\quad\therefore z = -1.6$

$\therefore \frac{K - 50}{5} = -1.6$ $\quad\therefore K - 50 = -8$ $\quad\therefore K = 50 - 8$ $\quad$ **$K = 42$**

Try to solve

6 If X is a normal random variable whose mean is μ, standard deviation is σ, $P(X < 19) = 0.7734$ and $P(X > 10) = 0.9332$, calculate the value of μ and σ.

Exercises 4 - 1

1 If Z is a standard normal random variable, find :

a. $P(0 \leq Z \leq 1.15)$, $P(0 \leq Z \leq 2.42)$

b. $P(-0.04 \leq Z \leq 0)$, $P(-1.63 \leq Z \leq 0)$

c. $P(-0.7 \leq Z \leq 0.7)$, $P(-1.65 \leq Z \leq 1.65)$

d. $P(-2.42 \leq Z \leq 1.67)$, $P(-1.73 \leq Z \leq 0.64)$

e. $P(0.74 \leq Z \leq 1.02)$, $P(1.4 \leq Z \leq 2.2)$

f. $P(-2.1 \leq Z \leq -0.92)$, $P(-1.5 \leq Z \leq -0.84)$

g. $P(Z \leq 1.44)$, $P(Z \leq 2.05)$

h. $P(Z \leq -1.14)$, $P(Z \leq -2.32)$

i. $P(Z \leq 0.65)$, $P(Z \leq 1.42)$

j. $P(Z \leq -0.45)$, $P(Z \leq -1.6)$

2 If Z is a standard normal random variable, find the value of the real number k which satisfies:

a. $P(0 \leq Z \leq K) = 0.3554$

b. $P(K \leq Z \leq 0) = 0.4120$

c. $P(-K \leq Z \leq K) = 0.2206$

d. $P(Z \leq K) = 0.9754$

e. $P(Z \leq K) = 0.1977$

f. $P(Z \geq K) = 0.0934$

g. $P(Z \geq K) = 0.9955$

h. $P(K \leq Z \leq 1.11) = 0.6660$

i. $P(K \leq Z \leq 2.22) = 0.2446$

j. $P(-1.7 \leq Z \leq K) = 0.3261$

3 Z is a standard normal random variable, if :

a. $P(Z \leq K) = 0.1736$ Find: $P(K \leq Z \leq 1.7)$

b) $P(Z \geqslant K) = 0.0207$ Find: $P(0.56 \leqslant Z \leqslant K)$

c) $P(Z \leqslant K) = 0.8944$ Find: $P(-0.7 \leqslant Z \leqslant K)$

d) $P(0.4 \leqslant Z \leqslant K) = 0.3110$ Find: $P(Z \leqslant K)$

e) $P(1.4 \leqslant Z \leqslant K) = 0.0770$ Find: $P(-1.4 \leqslant Z \leqslant K)$

f) $P(K \leqslant Z \leqslant 1.7) = 0.8586$ Find: $P(K \leqslant Z \leqslant 0.75)$

4) X is a normal random variable whose mean is μ, standard deviation is σ and

a) $P(X \leqslant 90) = 0.0668$, $\mu = 102$ Calculate σ

b) $P(X \geqslant 62) = 0.0548$, $\mu = 50$ Calculate σ

c) $P(X \geqslant 48) = 0.0228$, $\sigma = 4$ Calculate μ

d) $P(X > 68) = 0.1056$, $\sigma = 6.4$ Calculate μ

e) $P(X \geqslant 42) = 0.8944$, $\sigma = 6.4$ Calculate μ

f) $P(\mu - K\sigma \leqslant x \leqslant \mu + K\sigma) = 0.438$ Calculate K

g) $P(X \leqslant K) = 0.2119$, $\mu = 42$, $\sigma = 5$ Calculate K

h) $P(X \leqslant K) = 0.8413$, $\mu = 72$, $\sigma = 8$ Calculate K

i) $P(X > K) = 0.9772$, $\mu = 60$, $\sigma = 4$ Calculate K

5) Answer the following questions

a) If X is a normal random variable whose mean is 120, standard deviation is 10 and $p(X < K) = 0.9599$, find the value of k .

b) If X is a normal random variable whose mean is μ, standard deviation is $\sigma = 5$, find the value of μ which makes $p(X \leqslant 35) = 0.0228$

c) If X is a normal random variable whose mean is $\mu = 8$, standard deviation is $\sigma = 2$ and $P(K \geqslant X) = 0.1056$, find:

First : The value of k . **Second** : $P(X \leqslant 10)$

d) If X is a normal random variable whose mean is μ and standard deviation is σ , find $P(\mu - \frac{1}{4}\sigma \leqslant X \leqslant \mu + \frac{1}{2}\sigma)$

e If Z is a standard normal random variable, find the value of k which satisfies :

First : $P(Z > K) = 0.0281$

Second : $P(-1 < Z < K) = 0.7918$

f If X is a normal random variable whose mean is 18 and standard deviation is 2.5, find :

First: $P(X < 15)$

Second: $P(17 < X < 21)$

g If X is a normal random variable whose mean $\mu = 24$ and standard deviation $\sigma = 5$, find:

First : $P(X \geqslant 32.5)$

Second : $P(14 < X < 29)$

g If X is a normal random variable whose mean $\mu = 48$ and standard deviation $\sigma = 5$, find:

First : $P(43 < X < 59)$

Second : The value of K if $P(x > K) = 0.1841$.

h If X is a normal random variable whose mean $\mu = 17$ and standard deviation $\sigma = 2$, find:

First : $P(16 \leqslant X \leqslant 20)$

Second : $P(X > 15)$

i If X is a normal random variable whose mean is 32 and variance is 16, find :

First : $P(X < 25)$

Second : $P(28 < X < 35)$

j If X is a normal random variable whose mean $\mu = 8$ and standard deviation $\sigma = 2$, find:

First : $P(X \leqslant 10)$

Second : If $P(X \geqslant K) = 0.1056$, find the value of k .

Unit 4

4 - 2 Some practical applications of the normal distribution

You will learn

- Practical applications on the normal distribution

Key terms

- Normal distribution
- Normal curve
- Normal random variable
- Standard normal distribution

Introduction:

In the previous lesson, you became familiar with the normal distribution and its properties and you also knew the standard normal random variable and how to find it in terms of the mean and the standard deviation. You also knew how to calculate the probabilities of a random variable with a standard normal distribution using the statistical tables.

In this lesson, you are going to learn some different uses of the normal random variable for studying some phenomena it may express.

Example **Industry**

1. A machine in a factory produces cylinders of lengths follow a normal distribution whose mean is 56 cm and the standard deviation is 2 cm. The cylinder is only valid if its length ranges from 51 cm to 60 cm. A random sample has been randomly chosen out of 1000 cylinders. How many cylinders are expected to be valid?

Solution

Let X be a normal random variable expressing the cylinder length:

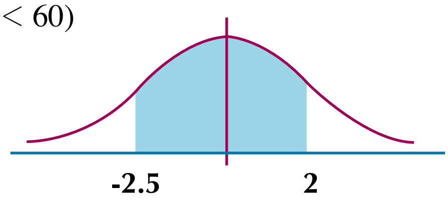
$\therefore$ The probability of (the cylinder is valid) = P (51 < X < 60)

$= P(\frac{51 - 56}{2} < Z < \frac{60 - 56}{2})$

$= P(-2.5 < Z < 2)$

$= P(-2.5 < Z \leqslant 0) + P(0 \leqslant Z < 2)$

$= 0.4938 + 0.4772 = 0.9710$



$\therefore$ The number of cylinders expected to be valid= $1000 \times 0.9710 = 971$ cylinders

Try to solve

1. Income: If the salaries of a 200- worker group at a factory follow a normal distribution whose mean is 175 LE and its standard deviation is 10 LE. How many workers do their salaries range from 170 LE to 180 LE?.

Materials
- Scientific calculator.

Example

(2) **Education:** If the marks of the students at a school are a normal random variable whose mean $\mu = 44$ and its standard deviation is σ where 22.66% of students have got more than 50 marks, find the value of σ.

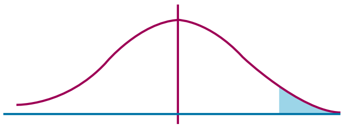
Solution

Let X be the normal random variable expressing the marks of the students

$\therefore P(X > 50) = \frac{22.26}{100}$

$\therefore P(Z > \frac{50 - 44}{\sigma}) = 0.2266$

$\therefore P(Z > K) = 0.2266$ where $K = \frac{6}{\sigma}$, $K > 0$



$\therefore P(0 \leqslant Z < K) = 0.5 - 0.2266 = 0.2734$

$\therefore K = 0.75$ $\therefore \frac{6}{\sigma} = 0.75$ $\therefore \sigma = \frac{6}{0.75} = 8$

Try to solve

(2) If the marks of the students in an exam follow a normal distribution whose mean is 60 and its standard deviation is 12 and a student has been randomly chosen, find the probability that the mark of that student is between 66 , 75 marks. If 15% of the excellent students have got an Excellent rank, find the least mark of the student getting an Excellent rank.

Example

(3) **Length:** If the lengths of the students at a high school follow a normal distribution whose mean $\mu = 160$ cm and its standard deviation $\sigma = 5$ cm, find the probability that the length of any student differs from μ not more than 8 cm .

Solution

Let X be a normal random variable expressing the lengths of the student and the length difference from $\mu = |x - \mu|$" i.e. the absolute difference between the length and the mean μ "

$\therefore P(|X - \mu| < 8) = P(|X - 160| < 8)$

$\therefore P(-8 < X - 160 < 8)$

$= P(152 < X < 168)$

$= P(\frac{152 - 160}{5} < Z < \frac{168 - 160}{5})$

$= P(-1.6 < Z < 1.6)$

$= 2 \times P(0 \leqslant Z < 1.6)$

$= 2 \times 0.4452 = 0.8904$

Remember

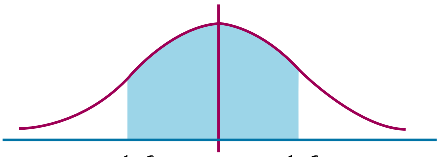
The expression:

$|x - a| < b$

is equivalent to the expression:

$-b < x - a < b$

I.e: $a - b < x < a + b$



Try to solve

3) Weights: If the weights of the students at a primary school follow a normal distribution whose mean is 30 kg and its standard deviation is 5 kg, calculate the percentage of the number of the students whose weights are more than 45 kg and so is the percentage of the number of the students whose weights range between 25 kg and 35 kg.

Example

4) Work: If the wages of the workers at a factory follow a normal distribution whose mean μ = 75 LE and its standard deviation σ = 10, find:

a) The percentage of the workers whose wages are more then 90 LE.

b) The percentage of the workers whose wages are less then 55 LE.

c) The percentage of the workers whose wages range between 60 LE and 80 LE.

Solution

a) $\because P(X > 90) = P(Z > \frac{90 - 75}{10})$

$= 0.5 - P(0 \leqslant Z \leqslant 1.5) = 0.5 - 0.4332 = 0.0668$

$\therefore$ The percentage of the workers whose wages are more then 90 LE = 6.68%

b) $\because P(X < 55) \quad = P(Z < \frac{55 - 75}{10}) \quad = P(Z < -2)$

$= 0.5 - P(0 \leqslant Z \leqslant 2) \quad = 0.5 - 0.4772 = 0.0228$

$\therefore$ The percentage of the workers whose wages are less then 55 LE = 2.28% of the total number

c) $\because P(60 \leqslant Z \leqslant 80) \quad = P(\frac{60 - 75}{10} \leqslant Z \leqslant \frac{80 - 75}{10})$

$= P(-1.5 \leqslant Z \leqslant 0.5) = P(0 \leqslant Z \leqslant 1.5) +$

$P(0 \leqslant Z \leqslant 0.5) = 0.1915 + 0.4332 = 0.6247$

$\therefore$ The percentage of the workers whose wages range between 60 LE and 80 LE = 62.47% of the total number of the workers at the factory.

Try to solve

4) Let the marks of an exam be a normal variable by expectation 76 and standard deviation 15 marks. By ordering the excellent students getting a mark higher than α mark, they represent 15% of the total students. By ordering the students getting marks lower than β mark, they represent 10% of the total students. Find:

a) The least mark α to consider the student belonging to excellent .

b) The failing mark β .

Exercises 4 - 2

1. If the salary of 1000 families in a city is a normal random variable whose mean is 170 LE, its standard deviation is 20 LE, and a family has been randomly chosen, find:
   - **a** The probability the salary is included between 160 LE and 200 LE .
   - **b** The number of families whose salaries are more than 150 LE.

2. If the weights of the students at a faculty follow a normal distribution whose mean is 68.5 kg and its standard deviation is 2.5 kg.
   - **a** calculate the percentage of the number of the students whose weights range between 67.5 kg and 71 kg.
   - **b** If the number of the students is 1000 students, calculate the number of the students whose weights are more than 71 kg .

3. A sample of 200 students at a school has been randomly taken. If their ages represent a normal random variable whose mean is 16.6 and its standard deviation is 1.2, find the number of the students whose ages are less than 16 years of this sample .

4. If the lengths of 2000 students at a faculty follow a normal distribution whose mean is 170 cm and its standard deviation is 8 cm, find the number of students whose lengths are less than 176 cm .

5. If the salary of 300 families represents a random variable x following the normal distribution by expectation μ = 500 LE and standard deviation σ = 20 L.E, find
   - **a** The number of the families getting a salary more than 350 LE .
   - **b** The maximal salary of 4% of the families getting the lowest salaries .

6. If the salary of 200 families is a random variable x following a normal distribution by expectation $\mu = 400$, standard deviation σ = 80 LE and a family has been randomly chosen, find:
   - **a** The probability the salary of the family is more than 500 LE at maximum
   - **b** The number of the families getting a salary 500 LE at maximum .

7. If the life time (in hours) of a type of batteries is a random variable following a normal distribution whose mean is 2000 hours and its standard deviation is 120 hours; What is the probability the battery would keep working for more than 1800 hours?

8) If the salary of a 500 - employee group follows a normal distribution whose mean is 180 LE and its standard deviation is 15 LE, find the numbers of employees whose salaries are less than 198 LE.

9) If the rainwater rise during February follows a normal distribution whose mean $\mu = 3$ cm and its variance $\sigma^2 = 4$ cm^2, find the probability the rainwater rise in February next year:

a) More than 1 cm

b) Between 3.5 cm and 4 cm

10) If the temperature in August follows a normal distribution whose mean $\mu = 35$ degrees and its standard deviation $\sigma = 5$ degrees, find the probability the temperature on a day during this month will be

a) Between 28 degrees and 38 degrees.

b) More than 39 degrees .

c) Between 26 degrees and 32 degrees.

11) 1000 adults have been applied to the enlisting management. If their lengths follow a normal distribution whose mean is 170 cm and its standard deviation is 10 cm, find the number of the adults:

a) Whose lengths are less than 190 cm

b) Unaccepted if the minimal required length is 155 cm

12) It is found that the lengths of a certain type of plants are distributed in regard to a normal distribution whose mean is 50 cm and its standard deviation is σ . If it is known that the lengths of 10.56% of these plants are less than 45 cm, find the variation of the lengths of these plants.

13) If the weights of the students in a faculty follow a normal distribution whose mean is 65 kg and its standard deviation is σ . If the weights of 33% if the students are more than 70 kg.

a) Find the value of σ

b) If the number of the students is 100 students, calculate the number of the students whose weights are less than 67.5 kg

14) If the weights of the students in a faculty follow a normal distribution whose mean is 68.5kg and its standard deviation is 2.5 kg:

a) Calculate the percentage of the students whose weights range between 67.5 kg and 71kg .

b) If the number of the students is 1000 students, calculate the number of the students whose weights are more than 71 kg.

15) If the marks of the students at a school are a normal random variable whose mean $\mu = 42$ and its standard variation is σ where 26.11% of the students got more than 50 marks, find the value of σ.

16 In a math exam, the marks of the students have a normal distribution whose mean is 70 and its standard deviation is 5, find the number of the students whose marks are more than 78 known that the number of the students did the exam are 100 students .

17 A factory produces cylinders of lengths follow a normal distribution whose mean is 56 cm and its standard deviation is 2 cm. The produced cylinders are valid if only their lengths are included between 51 cm and 60 cm. A random sample of 1000 cylinders has been taken. How many cylinders are expected to be valid?

18 If the radius lengths of the cylinders produced by a factory follow a normal distribution whose mean is 25 cm and its standard deviation is 20 cm. The cylinder is invalid if its radius length is less than 20 cm or more than 28 cm. If a cylinder has been randomly chosen, find the probability the cylinder is invalid.

19 If the weights of a group of the experimental animals follow a normal distribution whose mean is μ gm and its standard deviation is 10 gm and known that $P(x \geqslant 180) = 0.1587$, calculate the mean μ

20 If the marks of the students in an exam represent a random variable following a normal distribution whose mean is μ and its standard deviation is σ, find:

a The probability the students get a mark more than $(\mu - \sigma)$.

b The percentage of the students who get marks included between: $(\mu - 2\sigma)$ and $(\mu + 2\sigma)$.

21 It is found that the lengths of a certain type of plants are distributed in regard to a normal distribution whose mean is μ and its standard deviation is 4 . If it is known that the lengths of 10.56% of these plants are less than 45 cm, find the mean μ of the lengths of these plants

22 If the temperature in January follows a normal distribution whose mean is 16 degrees and its standard deviation is 4 degrees, find the probability the temperature on a day during this month will be:

a Between 14 degrees and 20 degrees.

b More than 15 degrees.

23 In a community, it is found that the intelligence follows a normal distribution whose mean is 104.6 and its standard deviation is 6.25,

a Find the ratio of the individuals whose intelligence ranges between 90 and 120

b Find the ratio of the individuals whose intelligence is more than 110.

Unit summary

Normal random variable

The range of the continuous random variable x is identified by the interval $]-\infty , \infty[$ and its probability density function is represented by a bell-like curve. The shape of the curve is identified by knowing two values: the mean μ and the standard deviation σ.

Some properties of the normal curve

1. It has one top and its two ends extends to $-\infty , \infty$.
2. It has a symmetrical axis passing through the top and intersecting the horizontal axis when $X = \mu$.
3. The area of the region placed under the normal curve and above the X axis equals 1.
4. From symmetry, we find that the straight line $X = \mu$ divides the area placed under the curve and above the X axis into two regions; the area of each region = 0.5.
5. The approximate area of the region under the curve and above the X axis can be calculated in regard to the following intervals:

- From $\mu - \sigma$ to $\mu + \sigma$ = 68.26 % of the total area .
- From $\mu - 2\sigma$ to $\mu + 2\sigma$ = 95.44 % of the total area .
- From $\mu - 3\sigma$ to $\mu + 3\sigma$ = 99.74 % of the total area .

The standard normal distribution

If the probability distribution of the random variable x is the normal distribution by a mean μ and standard deviation σ , then .$Z = \frac{x - \mu}{\sigma}$ is the standard normal distribution whose expectation is zero and its standard deviation = 1.

Properties of the density function of the standard normal distribution(Z):

1. The curve is located above the horizontal axis.
2. It is symmetrical in terms of the vertical axis.
3. The two ends of the curve extends infinitely without getting converged with the horizontal axis.
4. The area of the region under the curve and above the horizontal axis = 1.
5. From the symmetry, we find that the vertical axis divides the region under the curve and above the horizontal axis into two regions of area = 0.5 for each.
6. The approximate area of the region located off under the standard curve only and above any interval [a , b] by special tables can be calculated.

The tables of the area under the standard normal distribution curve:

To convert the normal distribution variable X into standard normal distribution variable Z, you use the relation:

$Z = \frac{x - \mu}{\sigma}$, from the table of the standard normal distribution attached at the end of the book, you can find the required area.

General exercises

1. Prove that for any normal distribution of a random variable x whose mean is μ and its standard deviation is σ, $P(x \geqslant \mu + 2\sigma) = 0.0228$

2. If x is a normal variable whose mean is μ and its variance is 64, find the value of μ if $P(x < 70) = 0.17$

3. Find σ if known $P(x < 37.25) = 0.0446$ and $\mu = 50$

4. If x is a normal random variable whose mean $\mu = 16$ and its standard deviation $\sigma = 4$, find the value of each :
   - a $P(x \leqslant 19)$
   - b $P(9 \leqslant x \leqslant 20)$
   - c $P(x \geqslant 10)$

5. If x is a normal random variable whose mean is μ, its standard deviation is σ, $P(x < 19) = 0.7734$ and $P(x > 10) = 0.9332$, find the value of μ and σ.

6. If x is a normal random variable whose mean is μ and its standard deviation is σ, find the value of each :
   - a $P(\mu - 2\sigma \leqslant x \leqslant \mu + 2\sigma)$
   - b $P(\mu - 1.35\sigma \leqslant x \leqslant \mu - 0.7\sigma)$ then calculate the value of each μ and σ

7. If x is a normal random variable whose mean is μ and its standard deviation is σ and:
   - a $P(x > 45) = 0.0228$, $\sigma = 10$ find μ
   - b $P(x \geqslant 15) = 0.8413$, $\sigma = 5$ find μ
   - c $P(x < 55) = 0.3085$, $\mu = 60$ find σ

8. If x is a normal variable whose mean is 50 and its standard deviation 10, find:
   - a $P(x > 70)$
   - b The value of k if $P(x < K) = 0.1587$

9. If x is a random variable of a normal distribution whose mean is 100 and its standard deviation is 4,
   - a If $P(x > a) = 0.5636$ find the value of a
   - b Find $P(x < 90)$
   - c Find $P(x > 108)$
   - d Find $P(95 < x < 105)$

10. If x is a normal random variable whose mean is μ, its standard deviation $8 = \sigma$ and $P(x \leqslant 40) = 0.1587$, find :
    - a The value of the mean
    - b $P(x > 50)$

11. If Z is a standard random variable and:
    - a $P(Z > K) = 0.1056$ find the value of k
    - b $P(-0.44 \leqslant Z \leqslant K) = 0.5588$ find the value of k

12. If x is a normal random variable whose mean is μ and its standard deviation is σ, find:

a. $P(x > \mu)$
b. $P(x < \mu)$
c. $P(\mu - \sigma \leqslant x \leqslant \mu + \sigma)$
d. $P(x > \mu + 2\sigma)$
e. $P(\mu - \sigma \leqslant x \leqslant \mu + 2\sigma)$
f. $P(\mu - 0.5\sigma \leqslant x \leqslant \mu + 0.5\sigma)$

13. If x is a normal variable whose mean is μ, its standard deviation is σ and $a > 0$, find a where

a. $P(\mu - a\sigma \leqslant x \leqslant \mu + a\sigma) = 0.6476$
b. $P(\mu - a\sigma \leqslant x \leqslant \mu + a\sigma) = 2\,P(0 \leqslant Z \leqslant 1.48)$

14. If x is a normal variable whose mean is 75 and its standard deviation is 15, find the value of k if $P(x > K) = 0.15$

15. If x is a normal random variable whose mean is μ and its standard deviation is 5, find the value of μ which makes $P(x \leqslant 45)$ equals:

a. 0.0228
b. 0.9332

16. **Weather:** If the temperature within a month follows a normal distribution whose mean is 20° and its standard deviation is $3\frac{1}{3}^\circ$, find the probability the temperature is between 21° and 25°

17. **Income:** The daily wages of a 1000- employee group follow a normal distribution whose mean is 400 LE, standard deviation is 80 LE and an employee is randomly chosen, find :

a. The probability the wage of an employee is less than 240 LE.
b. The percentage of the employees whose wages are more than 300 LE.
c. The number of employees whose wages are included between 260 LE and 340 LE.

18. **Wages:** If the wages of a 500-employee group follow a normal distribution whose mean is 60 and its standard deviation is 12, find the number of employees:

a. Whose wages are not greater than 54.
b. Whose wages are not less than 81.

19. **Exams:** If the marks of the students in an exam is a random variable following a normal distribution whose mean is μ and its standard deviation is σ, find:

a. The probability the students get a mark more than $(\mu - \sigma)$
b. The percentage of the students getting a mark included between $(\mu - 2\sigma)$ and $(\mu + 2\sigma)$.

For more activities and exercises www.sec3mathematics.com.eg

Accumulative test

1. A box contains 15 balls; 5 red balls are numbered from 1 to 5 and 10 black balls numbered from 6 to 15. If a ball has been randomly drawn .

First : Calculate the probability for each of the following two events:

A The event the drawn ball is red or carries an odd number.

B The event the drawn ball is black and carries an even number.

Second : Are the two events A and B mutually exclusive? Explain.

2. If X is a continuous random variable whose probability density function f(x), where:

$$f(x) = \begin{cases} K(2x-1) & \text{when } 1 \leqslant x \leqslant 3 \\ \text{zero} & \text{otherwise} \end{cases}$$

Find :

a The value of the constant k

b $P(X \geqslant 2)$

3. In a study conducted about the relation between the monthly income X and the monthly savings Y in LE for a 20- family sample, the gathered data have revealed the following:

$\Sigma X = 3000$ $\Sigma Y = 300$ $\Sigma X^2 = 800000$

$\Sigma Y^2 = 5500$ $\Sigma XY = 60000$

a Calculate the linear correlation coefficient between the income and the savings of the family .

b Find the regression line equation .

c Estimate the amount of money saved by a family whose monthly income is 2000 LE .

4. If X is a discrete random variable and you have the next function:

$f(x) = \frac{x^2 + K}{18}$ where $x = -1, 0, 2, 3$

a Find the value of k which makes f(x) a probability distribution function of the variable X.

b Calculate the mean of the variable x.

c Find $P(X \leqslant 2)$.

5. If A and B are two events of the sample space S for a random experiment and:

$P(A) = \frac{1}{2}$, $P(B) = x$ and $P(A \cup B)` = \frac{1}{3}$

First: Find the value of x in each of the following two cases:

a A and B are two mutually exclusive events .

b $A \subset B$

Second: If $x = \frac{1}{4}$ find $P(A \cap B)$.

6. **Length:** If the lengths of a group of students follow a normal distribution whose mean is μ and its standard deviation is 8 cm, find the value of μ if the standard length of a 180 cm tall student is 1.25 .

7. **Production :** If x and Y are two variables representing the production in a company where x represents the size of the production, Y represents the wages of employees in thousand LE, and the next data about 6 years are shown as follows .

Size of production	1000	2000	2500	4000	2300	2500
Wages	150	200	150	700	180	200

Calculate the value of the rank correlation coefficient between the size of the production and wages, then show its type.

8. **Bonus:** If the distribution of the bonuses of the workers at a factory follows a normal distribution whose mean μ = 75 LE, and its standard deviation σ = 10, find the percentage of the number of workers whose:
   - a) Bonuses are more than 90 LE.
   - b) Bonuses are less than 55 LE.
   - c) Bonuses are included between 60 LE and 80 LE.

9. **Weather:** If the temperature in January follows a normal distribution whose mean is 16 degrees and its standard deviation is 4 degrees, find the probability the temperature on a day within this month is.
   - a) Included between 14 degrees and 20 degrees
   - b) More than 15 degrees .

10. **Energy conservation:** At a factory for producing the saving electric lamps, it is noticed that the lifespan of the produced lamps per day follows a normal distribution whose mean is μ, its standard deviation is 20 days and the lifespan of 4.95% of the produced lamps is less than 100 days .
    - a) Find the value of μ
    - b) Within 10000 lamps of the produced lamps, estimate the number of the lamps whose lifespan ranges be 100 , 150 days.

Test 1

You may use the calculator and the area table:

Question 1:

(a) Complete the following sentences:

1. If A , B two events of sample space S for a random experiment where $P(B') = 0.6$, the value of $P(A' \cap B) + P(A) \times P(B \mid A) =$ ....

2. If Z is a standard normal variable where $P(K \leqslant Z \leqslant 1.5) = 0.03$, the value of K = ...

3. If A and B are two independent events of sample space S for a random experiment where $P(A) = 0.3$, $P(B) = 0.8$ then $P(A - B) =$ ....

4. If X is a discrete random variable, its expectation equals 5 , $\Sigma x_r^2 \cdot f(x_r) = 34$, then its standard deviation equals ....

5. If the regression line equation of y on x is $y = 0.2x + 3$ and the value table of y when $x = 5$ is 4.6 , than the error in the value of y = ...

(b) A and B are two events where $P(A) = 0.6$, $P(A' \cap B) = 0.2$, $P(A \cap B) = 0.3$ calculate:

a. $P(B \mid A)$

b. $P(A' \mid B')$

Question 2:

(a) The following table shows the marks of 6 students in a math exam (x) and statistics exam (y). calculate spearman's rank correlation coefficient between x and y and determine its type

x	Excellent	Pass	Good	Pass	Very good	Very good
y	Very good	Pass	Pass	Good	Good	Excellent

(b) If X is a normal random variable whose mean $\mu = 10$ and its standard deviation $\sigma = 2.5$

1. Find $P(X \leqslant 12.5)$

2. If $P(X \geqslant K) = 0.1056$, find the value of K.

Question 3:

(a) If x is a continuous random variable and :

$$f(x) = \begin{cases} \frac{x-1}{8} & \text{where} \quad 1 \leqslant x \leqslant 5 \\ \text{zero} & \text{otherwise} \end{cases}$$

1. Prove that f(x) the probability density function of the random variable X .
2. Calculate P (2 < x < 3)

(b) If the monthly income of 1000 families in a city is a normal random variable whose mean is 1700 LE and its standard deviation is 200 LE and a family is randomly chosen from these families, find the number of families whose monthly income is greater than 1500 LE.

Question 4 :

(a) If X is a discrete random variable whose range is { -2, -1 , 0 , 1, 2} and $P(X = r) = \frac{a + r}{15}$ for each r belongs to the range of x, find the value of a then find the standard deviation of the variable X.

(b) If : $\Sigma x = 49$, $\Sigma y = 45$, $\Sigma x^2 = 359$, $\Sigma y^2 = 303$, $\Sigma x\,y = 320$, $n = 7$

1. Calculate pearson's correlation coefficient between the values of x and y and determine its type.
2. Estimate the value of y when x=9 using the regression line

Test 2

You may use the calculator and the area table:

Question 1

(a) Choose the correct answer:

1. A regular die is rolled once, the probability of appearing the number 5 known that the number appeared is odd equals:

   a $\frac{1}{4}$ b $\frac{1}{3}$ c $\frac{1}{2}$ d $\frac{3}{4}$

2. If A and B are two events, $P(A \cap B) = 0.2$, $P(B) = 0.4$ then $P(A \mid B)$ equals

   a 0.5 b 0.06 c 0.14 d 0.1

3. The value of K in the following probability distribution is:

x_r	3	5	8
$f(x)_r$	k^2	$\frac{1}{4}$	$\frac{1}{2}$

   a $\frac{1}{4}$ b $\frac{1}{2}$ c $\frac{3}{4}$ d 1

4. If the marks of students in a classroom in statistics exam follow a normal distribution whose mean is 76, its standard deviation is 5 and Ahmed has got 66 marks in this exam, then his mark in the standard form is:

   a 3 b -2 c 1 d 2

5. The coefficient representing the strongest relation between two variables is:

(a) -0.58 (b) 0.48 (c) 0.68 (d) -0.78

(b) A box contains 9 balls identical in size and touch and they are numbered 0,1,2,......8. 2 balls are randomly drawn one after another without replacing. Calculate the probability:

1. The first ball carries an even number and the second carries an even number, too. (getting 2 even numbers).

2. The first ball carries an odd number and the second carries an even number.

Question 2:

(a) From the data of the following table :

x	150	180	150	120	120	100
y	120	120	100	80	80	100

Calculate spearman's rank correlation coefficient between x and y.

(b) If x is a discrete random variable whose probability distribution is as follows:

x_r	1	2	4	6
$f(x_r)$	0.2	a	0.4	0.1

Find the value of a, then calculate the mean and the standard deviation of the random variable x.

Question 3:

(a) If the salaries of a group of employees in a company follow a normal distribution whose mean is μ and its standard deviation σ = 250 LE and the percentage of the employees whose salaries are greater than 2150 LE is 97.72%, find the value of μ .

(b) If x is a continuous random variable, its probability density function is :

$$f(x) = \begin{cases} \frac{1}{8}(x + K) & \text{When} \quad 2 \leqslant x \leqslant 4 \\ \text{Zero} & \text{otherwise} \end{cases}$$

1. Find the value of K
2. Find $P(X < 3)$

Question 4 :

(a) If $\Sigma x = 40$, $\Sigma y = 30$, $\Sigma x^2 = 360$, $\Sigma y^2 = 200$, $\Sigma xy = 232$, $n = 5$ find :

1. Pearson's linear correction coefficient between x and y.
2. The regression line equation of y on x then estimate the value of y when x= 9

(b) If z is a standard random variable, find the value of K if : $P(z \geqslant K) = 0.1170$

Test 3

Question 1:

(a) Complete the following sentences:

1. If $P(B) = \frac{1}{2}$, $P(A \cup B) = \frac{5}{6}$ then $P(A' \mid B')$ equals ....

2. If X is a random variable whose range is {0, 1, 2, 3, 4} and $P(X = 0) = P(X = 4) = \frac{1}{16}$, $P(X = 1) = P(X = 3) = \frac{1}{4}$ then $P(X = 2)$ equals ....

3. If X is a continuous random variable and the probability density function is :

$$f(x) = \begin{cases} a & \text{when} \quad , -3 \leqslant x \leqslant 3 \\ \text{Zero} & \text{otherwise} \end{cases}$$

then a equals ....

4. If A and B are two independent events , P(A) = 0.3, P (B) = 0.6 then $P(A \cup B) = x$ then

$x =$ ....

5. If the lengths of a 1000 - people group follow a normal distribution whose mean is $176\frac{1}{2}$ and its standard deviation is 5 then the number of people whose lengths are greater than 185cm equals ....

(b) If A and B are two events of sample space for a random experiment S , prove that :
$P(B) = P(A) \times P(B \mid A) + P(A') \times P(B \mid A')$, then use this to calculate P (B)
If $P(A) = 0.6$, $P(B \mid A') = 0.8$, $P(B \mid A) = 0.3$

Question 2:

(a) If x is a continuous random variable and the probability density function is:

$$f(x) = \begin{cases} \frac{2x + 1}{24} & , 2 \leqslant x \leqslant 5 \\ \text{Zero} & \text{otherwise} \end{cases}$$

Calculate each of :

1. $P(3 \leqslant X \leqslant 5)$

2. $P(X \leqslant 4)$

(b) Find spearman's rank correlation coefficient between the two variables x and y from the data of the following table

x	10	16	15	17	18
y	5	7	8	6	9

Question 3:

(a) If x is a discrete random variable whose probability distribution is given by the function f where

$f(x) = \frac{x}{10}$, $x \in \{1, 2, 3, K\}$ find :

1. The value of K and write down the probability distribution of the random variable x
2. The expectation and variance of the random variable x

(b) If X x is a normal random variable whose mean $\mu = 50$, and its standard deviation is σ Find σ If $P(X \leqslant 37.25) = 0.0446$

Question 4:

(a) To study the relation between the required amount (y) in kg and the price (x) in LE of a certain product, we have the next data:

$\Sigma x = 25$, $\Sigma y = 30$, $\Sigma xy = 181$,

$\Sigma x^2 = 155$, $\Sigma y^2 = 249$, $n = 5$ find:

1. Pearson's correlation coefficient between x and y
2. The regression coefficient of the amount required by the price.

(b) If $P(B \mid A) = \frac{2}{3}$, $P(B \mid A') = \frac{5}{8}$, $P(A) = \frac{3}{4}$ find $P(A \cup B)$

Test 4

Question 1:

(a) Choose the correct answer:

1. A box contains 15 lamps out of them 5 lamps are defective. If two lamps are drawn one after another without replacing, the probability the two lamps are defective is:

   a $\frac{1}{3}$ b $\frac{2}{5}$ c $\frac{2}{7}$ d $\frac{2}{21}$

2. If A and B are two events of sample space for a random experiment S and $A \subset B$ then $P(B/A)$ equals

   a P(A) b P(B) c P(A -B) d P (S)

3. If all the points in a scatter diagram lie on a straight line, the correlation coefficient between the two variables equals:

   a ± 1 b 0 c $\frac{1}{2}$ d $\frac{3}{4}$

4. The value of R in the following probability distribution is :

x_r	–2	–1	1	2
$f(x_r)$	3R	$\frac{1}{4}$	2R	$\frac{1}{3}$

a $\frac{5}{7}$ b 1 c $\frac{1}{12}$ d $\frac{1}{3}$

5. If the probability distribution function of the random variable x is K

$$f(x) = \begin{cases} \frac{1}{3} & , \quad 1 \leqslant x \leqslant 4 \\ \text{Zero} & \quad \text{otherwise} \end{cases}$$

then P (X < 2) = ...

a $\frac{2}{3}$ b $\frac{1}{3}$ c $\frac{3}{4}$ d 1

(b) A and B are two events of sample space S for a random experiment , P (A') = $\frac{1}{4}$, P (B' | A) = $\frac{2}{5}$, calculate P (A ∩ B)

Question 2:

(a) If x is a discrete random variable whose probability distribution is as follows:

x_r	0	1	2	3	4
$f(x_r)$	0.25	0.2	0.1	0.3	0.15

Find the mean and the standard deviation of the random variable X.

(b) In an experiment of rolling a regular die two consecutive times and observing the number appeared on the upper face in each time. Calculate the probability of occurring the following events:

1. The appearance of two numbers their sum is greater than 8
2. The appearance of two numbers the absolute difference between them is less than 2 in a condition their sum is greater than 8

Question 3:

(a) The next table shows the grades of six students in math and physics exams:

Grades in physics	Pass	Good	Excellent	Weak	Very good	Good
Grades in math	Pass	Very good	Very good	Pass	Excellent	Weak

Calculate spearman's rank correlation coefficient and show its type.

(b) If x is a continuous random variable whose probability density function is:

$$f(x) = \begin{cases} \frac{1}{4}(x + a) \text{ where } & 0 \leqslant x \leqslant 2 \\ \text{Zero} & \text{otherwise} \end{cases}$$

Find the value of a, then calculate $P(\frac{1}{2} \leqslant x \leqslant \frac{3}{2})$

Question 4:

(a) If the mean of a random variable equals 150 and the coefficient of variation equals 2%, find the variance of this random variable.

(b) In a study to find the relation between the weight x in kg and the length y in cm for six people, it is found that:

$\Sigma x = 374$, $\Sigma y = 913$, $\Sigma x^2 = 24364$,

$\Sigma y^2 = 193624$, $\Sigma xy = 52260$ find:

1. Pearson's linear correlation coefficient between x and y
2. The regression line equation of y on x.

Test 5

Question 1:

(a) Choose the correct answer:

1. If P(A') = 0.3 , P(b) = 0.4 , P (A ∩ B) = 0.2 then P (A | B') =

   a $\frac{1}{2}$ b $\frac{5}{6}$ c 1 d $\frac{3}{4}$

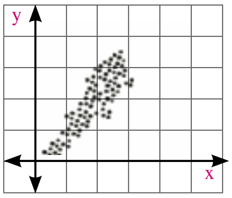
2. The coefficient representing the strongest correlation between the two variables is:

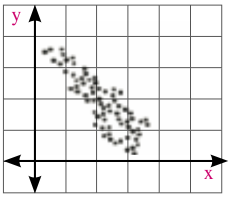
   a -0.2 b 0.1 c -0.8 d 0.7

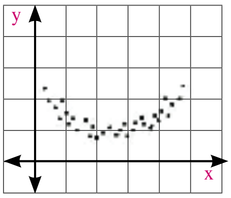
3. If x is a random variable whose range is { 1, 2, 3, 5} , and P (x = 1) = 2 P (x = 2) = $\frac{1}{4}$, P (x = 3) = $\frac{7}{16}$ then P (x = 5) equals:

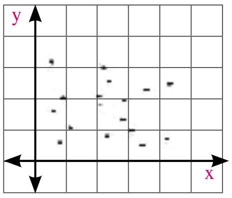
   a $\frac{3}{8}$ b $\frac{3}{16}$ c $\frac{3}{4}$ d $\frac{11}{16}$

4. The scatter diagram representing a direct correlation is figure

   a 

   b 

   c 

   d 

5. If there is a relation between two variables x and y $\Sigma X_r \cdot f(x_r) = 4$, $\Sigma x_r^2 \cdot f(x_r) = 25$ then the coefficient of variation equals:

a) 16٪ b) 75٪ c) 64٪ d) 15.6٪

(b) If A and B are two independent events of sample space S for a random experiment , $P(A) = 2P(B) = x$, $P(A \cup B) = \frac{7}{9}$ find the value of x.

Question 2:

(a) If x is a continuous random variable whose probability density function is:

$$f(x) = \begin{cases} \frac{1}{32}(x+4) \text{ where } & -K \leqslant x \leqslant K \\ \text{Zero} & \text{otherwise} \end{cases}$$

Find:

1. The value of K 2. $P(x \leqslant 0)$ 3. $P(-2 \leqslant x \leqslant 2)$

(b) The following table shows the grades of eight university students in math and physics exams:

Grades in math(x)	Excellent	Good	Very good	Very good	Weak	Excellent	Pass	Very good
Grades in physics(y)	Very good	Very good	Good	Excellent	Pass	Excellent	Pass	Excellent

Find spearman's rank correlation coefficient among the marks of math and physics and determine its type.

Question 3:

4. If the weights of the students at a university follow a normal distribution whose mean is 55kg, its standard deviation is σ and the weights of 33 % of the students are greater than 66kg, find:

1. The standard deviation
2. If the number of the students is 10000, calculate the number of the students whose weights are less than 60 kg.

(b) If x is a discrete random variable whose mean $\mu = 3$ and its probability distribution is as follows:

x_r	zero	K	3	4
$f(x_r)$	m	$\frac{1}{6}$	4m	5m

Find:

1. The value for each of M and K

2. The standard deviation and the coefficient of variation of variable x.

Question 4:

(a) a box contains five identical cards numbered from 1 to 5. Two cards are drawn one after another with replacing. Find the probability:

1. The sum of the two numbers on the two cards is a prime number.
2. The product of the two numbers is less than seven if their sum is a prime number.

(b) In a study to show the relation between two variables x and y, we have got the following results:

$n = 10$, $\Sigma x = 35$, $\Sigma y = 60$, $\Sigma xy = 187$

$\Sigma x^2 = 134$, $\Sigma y^2 = 406$ find:

1. K
2. The regression line equation of y on x.
3. Pearson's linear correlation coefficient between x and y, then determine its type.

Test 6

Question 1:

(a) Choose the correct answer:

1. The strongest correlation coefficient of the following is:

   a) 0.7 b) 1.2 c) -0.9 d) -0.3

2. If x is a random variable whose range is {1, 2, 3} then the function representing the probability distribution function is:

   a) $f(x) = \frac{x+2}{8}$ b) $f(x) = \frac{x+1}{3}$ c) $f(x) = \frac{x}{6}$ d) $f(x) = \frac{2x+3}{6}$

3. In an experiment of rolling a regular die once, the probability of appearing an odd number if a number less than 4 appeared is:

   a) $\frac{1}{4}$ b) $\frac{2}{3}$ c) $\frac{3}{4}$ d) $\frac{1}{2}$

4. If A and B are two independent events , P(A) = 0.5, P(B) = 0.6 then $P(A \cup B) =$ ....

   a) 0.3 b) 1.1 c) 0.8 d) 0.1

5. If x is a normal variable whose mean $\mu = 6$ and its standard deviation $\sigma = 3$ then the variable subjected to a standard normal distribution is:

   a) $\frac{6-X}{3}$ b) $\frac{X-3}{6}$ c) $\frac{X-6}{3}$ d) $\frac{3-X}{6}$

(b) If A and B are two events of a sample space for a random experiment , P (A') =0.6 , P(B) = 0.3 , P (A ∪ B) =0.9 calculate P (A | B')

Question 2:

(a) The next table shows the grades of six students in math(x) and statistics (y) exams:

x	Good	Very good	Weak	Excellent	Pass	Weak
y	Pass	Weak	Good	Pass	Very good	Excellent

Calculate spearman's rank correlation coefficient and show its type .

(b) If x is a continuous random variable

$$f(x) = \begin{cases} \frac{1}{18}(x+2) & \text{where } -2 \leqslant x \leqslant 4 \\ \text{Zero} & \text{otherwise} \end{cases}$$

First: prove that f(x) is the probability density function of the random variable x:

Second: Find $P(0 \leqslant x \leqslant 2)$

Question 3:

(a) If x is a normal random variable whose mean is μ and its standard deviation is σ calculate:

1. $P(\mu \leqslant x \leqslant \mu + \sigma)$
2. $P(\mu - \frac{7}{4}\sigma \leqslant x \leqslant \mu + \frac{7}{4}\sigma)$
3. $P(x - \mu \leqslant 1.8\sigma)$

(b) If x is a discrete random variable whose probability distribution is as follows:

x_r	−1	0	1	2	4
$f(x_r)$	2P	P	3P	2P	P

Find the value of P, then calculate the mean and variance of the random variable x.

Question 4:

In a study to show the relation between two variables x and y, If know that:

$\Sigma x = 540$, $\Sigma y = 460$, $\Sigma x^2 = 37112$

$\Sigma y^2 = 28252$, $\Sigma xy = 30782$, $n = 8$

First: Calculate Pearson's linear correlation coefficient between x and y.

Second: Estimate the value of y when x=30 using the regression line equation.

Test 7

Question 1:

(a) Complete the following:

1. If the two points (6,5) and (3,8) lie on the regression line y on x, then the type of the correlation between x and y is ....

2. If x is a random variable whose range is { -1, 0, K} , and its probability distribution function is $f(x) = \frac{x+2}{4}$ then K = ...

3. If $P(A \cap B) = \frac{3}{8}$, $P(A') = \frac{1}{4}$ then $P(B \mid A)$ = ...

4. If A and B are two independent events, P(A) = 0.4 , $P(A \cup B) = 0.58$, then P (B) = ...

5. If x is a normal variable whose mean $\mu = 4$ and its variance = 25 , then $P(x \geqslant 14)$ = ...

(b) If x is a random variable whose mean $\mu = 55$ and its standard deviation is σ find the variance satisfying: $P(x \leqslant 45) = 0.228$.

Question 2:

(a) If x is a discrete random variable whose range is { -3, -2, 1, 2} , $P(x = -3) = P(x = -2) = \frac{1}{8}$, $P(x = 2) = \frac{1}{2}$, Calculate:

First: P(x = 1) **second**: the expectation, variance of the random variable x.

(b) The following table shows the grades of six students in math and chemistry exams

Grades in math	Very good	Weak	Pass	Good	Excellent	Pass
grades in chemistry	Pass	Good	Very good	Pass	Weak	Excellent

Calculate spearman's rank correlation coefficient and determine its type.

Question 3:

(a) If x is a continuous random variable whose probability density function is:

$$f(x) = \begin{cases} \frac{1}{24}x + K & \text{where } 1 \leqslant x \leqslant 5 \\ \text{Zero} & \text{otherwise} \end{cases}$$

Calculate the value of K, then find $P(2 \leqslant X \leqslant 4)$

(b) If it is known that the probability the weather is going to be rainy is 0.24, the probability it is stormy is 0.36 and the probability it is rainy and stormy is 0.14, calculate the probability of the following events:

First: the weather is either rainy or stormy.

Second: the weather is rainy where it is not stormy.

Question 4:

(a) In a study to show the relation between two variables x and y, if known that:
$\Sigma x = 477, \Sigma y = 222, \Sigma xy = 15184, \Sigma x^2 = 32593, n = 7$
find the regression line equation y on x then estimate the value of y when x = 100.

(b) If x is a normal random variable whose mean is μ, its standard deviation $\sigma = 8$, $P(x > 64) = 0.0668$ calculate the mean μ.

Test 8

Question 1:

(a) Choose the correct answer

1. If the regression line equation y on x is y= 0.2x+3, then the expected value of y when x = 5 is:

   a) 0.2   b) 3.2   c) 3   d) 4

2. If x is a random variable and the expectation equals 3 , $\Sigma x_r^2 \cdot f(x_r) = 14.5$ then standard deviation equals:

   a) 11.5   b) 5.5   c) 4.8   d) 2.35

3. If $P(A) = \frac{1}{2}$, $P(A - B) = \frac{3}{8}$ then $P(B' \mid A) =$:

   a) $\frac{3}{8}$ b) $\frac{3}{4}$ c) $\frac{9}{32}$ d) $\frac{3}{16}$

4. If A , B two independent events , P(A) = 0.6, P(B) = 0.3 then P (B - A) equals

   a) 0.9   b) 0.3   c) 0.18   d) 0.12

5. If Z is a standard normal variable, Then : $P(Z \geqslant 1.5)$ equals to the nearest two decimals :

   a) 2.23   b) 1.51   c) 0.07   d) 1.21

(b) In an experiment to roll two different dices once, calculate the probability the sum of the two numbers appeared on the two upper faces is an odd number known that the number appeared on the first upper face is 4

Question 2:

(a) If X is a continuous random variable whose probability density function is :

$$f(x) = \begin{cases} \frac{1}{12}(2x+1) & \text{where } 0 \leqslant x \leqslant 3 \\ \text{Zero} & \text{otherwise} \end{cases}$$

Calculate: **First**: $P(x \leqslant 2)$ **Second**: $P(1 \leqslant x \leqslant 2)$.

(b) The following table shows the marks of six students in math and statistics exams

Marks in statistics	22	25	19	24	25	13
Marks in math	45	35	40	28	40	25

Calculate spearman's correlation coefficient between the marks of math and statistics then show its type.

Question 3 :

(a) In a math exam, the marks of the students have been followed a normal distribution whose mean is 70 and its standard deviation is 5. Calculate the number of students whose marks are probably greater than 78 if it is known that the number of students applied to this exam is 10000.

(b) If x is a random variable whose probability distributions is as follows:

x_r	-2	-1	0	3	4
$f(x_r)$	$\frac{1}{4}a$	$\frac{1}{16}$	$\frac{3}{16}$	$\frac{1}{8}a$	$\frac{3}{8}$

Find the value of a, then find the mean and variance of the variable x.

Question 4:

(a) In a study to show the relation between two variables x and y, we have got the following results:

$n = 7$, $\Sigma x = 147$, $\Sigma y = 99$, $\Sigma xy = 2123$, $\Sigma x^2 = 3430$

① Find the regression line equation y on x ② Estimate the value of y when x= 20

(b) If A and B are two independent events $P(A) = 0.6$, $P(A - B) = 0.36$ calculate $P(A \cup B)$

Test 9

Question 1:

(a) Complete the following:

① If $P(A`) = \frac{2}{5}$ and $P(B|A) = \frac{1}{2}$ then $P(A \cap B) =$ ....

② If the probability distribution of the random variable x is:

x_r	-1	0	2
$f(x_r)$	$\frac{1}{6}$	$\frac{1}{3}$	$\frac{1}{2}$

Then the expectation equals ....

3 If A and B are two independent events of S , P (B') = 0.4 and P (A ∪ B) = 0.8 then P (A) = ...

4 In the regression line equation y on x: y = bx + a. If the coefficient of x is less than zero, then the correlation between the two variables x and y is ....

5 If x is a normal random variable whose mean is 75 and its standard deviation is 4, then P (X < 85) = ...

(b) A 42 -student classroom; 28 study English, 21 study Italian and 7 study both languages. A student is randomly chosen. Calculate the probability the student chosen studies:

First: Both languages **Second**: English if he (she) is already studying Italian

Question 2:

(a) From the data of the following table

x	9	12	11	14	10	12
y	15	20	19	23	17	18

First: Calculate spearman's correlation coefficient and show its type.

Second: Use the regression line to estimate the value of y and calculate the error when x = 11

(b) If A and B are two events of a sample space of a random experiment P(B'|A) = 0.4 ,

P (A |B) = 0.3 and P(A) + P (B) = 0.36 find P(A) and P (A ∪ B)

Question 3:

(a) If x is a continuous random variable

$$f(x) = \begin{cases} \frac{1}{8}x & \text{where } 3 \leqslant x \leqslant 5 \\ \text{Zero} & \text{otherwise} \end{cases}$$

First: Prove that f(x) is a density function of the random variable x

Second: find P (x > 4)

(b) If x is a normal random variable whose mean μ = 15 and its standard deviation σ = 5, find:

1 P (12 < x < 17)

2 The value of K where P (X < K) = 0.3446

Question 4:

(a) If x is a discrete random variable whose range is { 0, 1, 3, 5}, $P(X = 0) = \frac{1}{2}$, $P(x = 1) = \frac{1}{6}$ and $P(x = 3) = \frac{1}{4}$ find:

First: The probability distribution of the variable x

Second: The mean and the coefficient of variation of the variable x

(b) A box contains 7 white balls, 8 red balls and 5 black balls. Two balls are drawn respectively without replacing, calculate the probability:

First: The second ball is white if the first is white.

Second: The first ball is red if the second is red.

Test 10

Question 1:

(a) Choose the correct answer:

1. If x is a random variable whose range is {0, 1, 2, 3}, $P(x=0)=\frac{1}{8}$ $P(x=2)=\frac{1}{2}$ and $P(x=3)=\frac{1}{4}$ then P (x = 1) equals ....

   a) $\frac{1}{4}$ b) $\frac{1}{8}$ c) $\frac{1}{2}$ d) 1

2. If P (A -B) = 0.04 , $P(A \cap B) = 0.1$ then $P(B' \mid A)$ equals:

   a) 0.3 b) 0.5 c) 0.04 d) 0.8

3. If A and B are two mutually exclusive events , P (A) = 0.3 and P (B') = 0.4 then $P(A \cup B)$ equals ....

   a) 0.3 b) 0.7 c) 0.9 d) 0.6

4. If x is a random variable whose mean $\mu = 45$ and its standard deviation $\sigma = 5$, then $P(x \geqslant 55)$ equals ....

   a) 0.4772 b) 0.9772 c) 0.0228 d) 0.2386

5. If the regression line equation y on x is : y = 3-x, then the correlation between the values of x and the values of y is:

   a) nihilistic b) direct c) inverse d) perfect

(b) If the marks of students at a school follow a normal distribution whose mean $\mu = 42$ and its standard deviation is σ where 26.11% of students got more than 50 marks, find σ.

Question 2:

(a) Find spearman's correlation coefficient from the data of the following table and show its type:

x	1	2	3	4	5	6
y	35	12	17	25	12	7

(b) If A and B are two events of a sample space of a random experiment P (A ∪ B) = 0.8 , P(B) = 0.4 and P (A´) = 0.3 calculate P (B' | A)

Question 3 :

(a) x is a continuous random variable whose probability density function is

$$f(x) = \begin{cases} \frac{1}{6}x + K & \text{where } 0 \leqslant x \leqslant 3 \\ \text{Zero} & \text{otherwise} \end{cases}$$

Find the value of K, then find P (1.5 ⩽ X ⩽ 2.5)

(b) If A and B are two independent events of a sample space of a random experiment where A and B are two independent events, then P(B) = 0.4 and P(A ∩ B) = 0.24 calculate P (A ∪ B') and P (A ∪ B)

Question 4:

(a) Two regular dices; a number of 1.3 and 5 is written on each two opposite faces on the first die and a number of 2.4 and 6 is written on each two opposite faces on the second die. If the two dices are rolled and the random variable x expresses the sum of the two numbers appeared, find the probability distribution of the variable x and calculate the expectation and the coefficient of variation.

(b) The following data represent the expenditure y and the income x in LE per day of a sample.:

x	50	60	45	70	65	90
y	45	35	35	50	55	70

First: Estimate the expenditure if the income is 63

Second: Calculate the error when x = 45

Table of areas under the standard normal distribution curve

Z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.0000	0.0040	0.0080	0.0120	0.0160	0.0199	0.0239	0.0279	0.0319	0.0359
0.1	0.0398	0.0438	0.0478	0.0517	0.0557	0.0596	0.0636	0.0675	0.0714	0.0753
0.2	0.0793	0.0832	0.0871	0.0910	0.0948	0.0987	0.1026	0.1064	0.1103	0.1141
0.3	0.1179	0.1217	0.1255	0.1293	0.1331	0.1368	0.1406	0.1443	0.1480	0.1517
0.4	0.1554	0.1591	0.1628	0.1664	0.1700	0.1736	0.1772	0.1808	0.1844	0.1879
0.5	0.1915	0.1950	0. 1985	0.2019	0.2054	0.2088	0.2123	0.2157	0.2160	0.2224
0.6	0.2259	0.2291	0.2324	0.2357	0.2389	0.2422	0.2454	0.2486	0.2517	0.2549
0.7	0.2580	0.2611	0.2642	0.2673	0.2704	0.2734	0.2764	0.2794	0.2823	0.2852
0.8	2881	0.2910	0.2939	0.2967	0.2995	0.3023	0.3051	0.3078	0.3106	0.3133
0.9	0.3159	0.3186	0.3212	0.3238	0.3264	0.3289	0.3315	0.3340	0.3365	0.3389
1.0	0.3413	0.3438	0.3461	0.3485	0.3508	0.3531	0.3554	0.3577	0.3599	0.3621
1.1	0.3643	0.3665	0.3686	0.3708	0.3729	0.3749	0.3770	0.3790	0.3815	0.3830
1.2	0.3849	0.3869	0.3888	0.3907	0.3925	0.3944	0.3962	0.3980	0.3997	0.4015
1.3	0.4032	0.4049	0.4066	0.4082	0.4099	0.4115	0.4131	0.4147	0.4162	0.4177
1.4	0.4192	0.4207	0.4222	0.4236	0.4251	0.4265	0.4279	0.4292	0.4306	0.4319
1.5	0.4332	0.4345	0.4357	0.4370	0.4382	0.4394	0.4406	0.4418	0.4429	0.4441
1.6	0.4452	0.4463	0.4474	0.4484	0.4495	0.4505	0.4515	0.4525	0.4535	0.4545
1.7	0.4554	0.4564	0.4573	0.4582	0.4591	0.4599	0.4608	0.4616	0.4625	0.4633
1.8	0.4641	0.4649	0.4656	0.4664	0.4671	0.4678	0.4686	0.4693	0.4699	0.4706
1.9	0.4713	0.4719	0.4726	0.4732	0.4738	0.4744	0.4750	0.4756	0.4761	0.4767
2.0	0.4772	0.4778	0.4783	0.4788	0.4793	0.4798	0.4803	0.4808	0.4812	0.4817
2.1	0.4821	0.4826	0.4830	0.4834	0.4838	0.4842	0.4846	0.4850	0.4854	0.4857
2.2	0.4861	0.4864	0.4868	0.4871	0.4875	0.4878	0.4881	0.4884	0.4887	0.4890
2.3	0.4893	0.4896	0.4898	0.4901	0.4904	0.4906	0.4909	0.4911	0.4913	0.4916
2.4	0.4918	0.4920	0.4922	0.4925	0.4927	0.4929	0.4931	0.4932	0.4934	0.4936
2.5	0.4938	0.4940	0.4941	0.4943	0.4945	0.4946	0.4948	0.4949	0.4951	0.4952
2.6	0.4953	0.4955	0.4956	0.4957	0.4959	0.4960	0.4961	0.4962	0.4963	0.4964
2.7	0.4965	0.4966	0.4967	0.4968	0.4969	0.4970	0.4971	0.4972	0.4973	0.4974
2.8	0.4974	0.4975	0.4976	0.4977	0.4977	0.4978	0.4979	0.4979	0.4980	0.4981
2.9	0.4981	0.4982	0.4982	0.4983	0.4984	0.4984	0.4985	0.4985	0.4986	0.4986
3.0	0.4987	0.4987	0.4987	0.4988	0.4988	0.4989	0.4989	0.4989	0.4990	0.4990
3.1	0.4990	0.4991	0.4991	0.4991	0.4992	0.4992	0.4992	0.4992	0.4993	0.4993
3.2	0.4993	0.4993	0.4994	0.4994	0.4994	0.4994	0.4994	0.4995	0.4995	0.4995
3.3	0.4995	0.4995	0.4995	0.4996	0.4996	0.4996	0.4996	0.4996	0.4996	0.4997
3.4	0.4997	0.4997	0.4997	0.4997	0.4997	0.4997	0.4997	0.4997	0.4997	0.4998
3.5	0.4998	0.4998	0.4998	0.4998	0.4998	0.4998	0.4998	0.4998	0.4998	0.4998

Answer of the Exercises

Unit 1: Correlation and regression

Answers of exercises (1 - 1)

1. d
2. d
3. b
4. c
5. d
6. 0.9 $r \simeq$ 0.947 direct
7. 0.8857
8. - 0.985 inverse
9. 0.3375 direct
10. - 0.41 inverse
11. - 0.74 inverse
12. $\simeq$ - 0.91 inverse
13. 0.7285714
14. 0.166 direct
15. 0.8545 direct
16. 0.8661 direct

Answers of exercises (1 - 2)

1. b
2. b
3. c
4. a
5. d
6. d
7. y = 0.7 + 0.723 x , $\simeq$ 9.376
8. a) $\simeq$ 8.8592   b) = 1.2868
9. a) $\simeq$ 0.69999   b) y = 2.8x - 12.4
10. a) y = 36.85 - 0.6333x
    b) - 0.9237 inverse
11. a) - 0.915227 inverse
    b) y = 94.493 - 9.5628 x
12. a) 0.76856 direct.
    b) y = 8.4825 + 0.40381 x
    c) y $\simeq$ 28.67 L.E
    d) 28   error = 3.37
13. a) y =1.35 x - 0.375
    b) $\simeq$ 944.63.

Answers of general exercises

1. a
2. c
3. d
4. d
5. b
6. a
7. r $\simeq$ 0.47 direct
8. r $\simeq$ 0.931 direct
9. r $\simeq$ 0.81 direct.
10. r $\simeq$ -0.41 reverse
11. r $\simeq$ 0.89337
12. a) r $\simeq$ 0.55296
    b) y = 2.8223 + 0.3249 x   y $\simeq$ 5.096
13. r $\simeq$ 0.93 , $y^8 \simeq$ 11
14. a) y = 11.5483 - .2053 x
    b) y $\simeq$ 3.334858
    c) error = |3.334 - 4.77 | $\simeq$ 1.439

Answer of the accumulative test

1. a
2. c
3. c
4. a
5. c
6. b
7. If the sign is positive, the correlation is direct and if it is negative, the correlation is inverse.
8. a) r $\simeq$ - 0.5211
   b) y^8= 135.17 - x ,$y^8 \simeq$ 73.17
   c) error = | 72 - 69.17| = 2.83
9. a) r $\simeq$ - 0.9426
   b) y^8 = 10.38 - 0.154 x
   c) 4.22
   d) error $\simeq$ 1.528

Unit 2: Conditional probability

Answers of exercises (2 - 1)

1. b
2. a
3. $\frac{1}{3}$
4. $\frac{1}{2}$
5. a
6. **First:** 0.21   **Second:** 0.89   **Third:** 0.525   **Fourth:** $\frac{19}{30}$
7. $\frac{3}{5}$
8. **First:** $\frac{2}{5}$ **Second:** $\frac{29}{35}$
9. $\frac{2}{3}$
10. **First:** $\frac{1}{6}$ **Second:** $\frac{1}{2}$
11. 0.42

1. First: $\frac{11}{15}$ Second: $\frac{3}{5}$ Third: $\frac{1}{3}$
2. First: $\frac{1}{11}$ Second: $\frac{1}{4}$ Third: $\frac{4}{9}$
3. $\frac{1}{4}$
4. First: $\frac{1}{10}$ Second: $\frac{7}{16}$
5. $\frac{3}{4}$
6. First: $\frac{4}{11}$ Second: $\frac{5}{33}$ Third: $\frac{35}{132}$
7. First: $\frac{26}{35}$ Second: $\frac{55}{113}$
8. First: $\frac{5}{7}$ Second: $\frac{1}{2}$ Third: $\frac{13}{30}$
9. First: $\frac{3}{4}$ Second: $\frac{1}{2}$
   Third: $\frac{3}{16}$ Fourth: $\frac{13}{14}$
10. Activity: First: $\frac{5}{6}$ Second: $\frac{2}{3}$

Answers of exercises (2 - 2)

1. a , g , z are independent events.
   b, d, h, o are dependent events
2. c
3. b
4. d
5. $\frac{1}{12}$
6. $\frac{1}{16}$
7. a , b are two dependent events
8. First : 0.2 Second: $\frac{2}{7}$
9. $\frac{1}{4}$
10. $\frac{1}{15}$
11. a $\frac{3}{95}$ b $\frac{9}{190}$
    c $\frac{15}{190}$ d $\frac{2}{95}$
12. First: a 0.28 b 0.82
    c 0.54 d 0.72
    Second: $\frac{2}{5}$

Answers of general exercises

1. c
2. b
3. d
4. c
5. a
6. $\frac{1}{10}$
7. $\frac{2}{5}$
8. $P(A \cap B) = P(A) \times P(B) = \frac{3}{8}$
9. a $\frac{5}{96}$ b $\frac{15}{96}$ c $\frac{1}{16}$ d $\frac{3}{16}$
10. $\frac{1}{2}$ $\frac{1}{8}$ 1
11. a $\frac{1}{6}$ b $\frac{5}{9}$ c $\frac{4}{9}$
12. a 0.48 b 0.92 c 0.8
13. a $\frac{9}{65}$ b $\frac{24}{65}$ c $\frac{16}{65}$ d $\frac{16}{65}$
14. a $\frac{1}{5}$ b $\frac{1}{4}$
15. a $\frac{16}{25}$ b $\frac{3}{4}$ c $\frac{1}{5}$
16. a $\frac{1}{4}$ b $\frac{2}{9}$

Answer of the accumulative test

1. Zero
2. $\frac{1}{6}$
3. $\frac{2}{5}$
4. 1
5. b
6. b
7. c
8. $-\frac{45}{56}$
9. a $\frac{14}{15}$ b $\frac{1}{3}$ c $\frac{1}{5}$
10. a $\frac{33}{95}$ b $\frac{21}{95}$ c $\frac{14}{95}$

Unit 3: Random variables and probability distribution

Answers of exercises (3 - 1)

1. c
2. a
3. b
4. a
5. d
6. d
7. a $\frac{1}{8}$ b $\frac{1}{10}$ g $-\frac{4}{5}$
8. 0.1
9. m = 1

x_r	- 2	0	2	4
$f(x_r)$	$\frac{1}{5}$	$\frac{2}{5}$	$\frac{1}{5}$	$\frac{1}{5}$

10. a = 3

x_r	1	2	3	4
$f(x_r)$	$\frac{9}{54}$	$\frac{12}{54}$	$\frac{15}{54}$	$\frac{18}{54}$

11. k = 5

x_r	1	2	3	4
$f(x_r)$	$\frac{8}{50}$	$\frac{11}{50}$	$\frac{14}{50}$	$\frac{17}{50}$

12.

x_r	-3	-1	1	3
$f(x_r)$	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{3}{8}$	$\frac{1}{8}$

13.

x_r	6	7	8	9	10
$f(x_r)$	$\frac{1}{9}$	$\frac{2}{9}$	$\frac{3}{9}$	$\frac{2}{9}$	$\frac{1}{9}$

14.

x_r	1	2	3	4	5	6
$f(x_r)$	$\frac{11}{36}$	$\frac{9}{36}$	$\frac{7}{36}$	$\frac{5}{36}$	$\frac{3}{36}$	$\frac{1}{36}$

1

x_r	1	1.5	2	2.5	3	3.5	4
$f(x_r)$	$\frac{1}{16}$	$\frac{2}{16}$	$\frac{3}{16}$	$\frac{4}{16}$	$\frac{3}{16}$	$\frac{2}{16}$	$\frac{1}{16}$

2 Range of x = { 0, 1, 2, 3}

x_r	0	1	2	3
$f(x_r)$	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{3}{8}$	$\frac{1}{8}$

3 Activity :

x_r	2	3	4	5	6	7	8	9	10	11	12
$f(x_r)$	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$

First: (a) $\frac{1}{2}$ (b) $\frac{1}{2}$

(c) $\frac{5}{6}$ (d) $\frac{15}{36}$

Second: The expected value is 7

Answers of exercises (3 - 2)

1 b **2** b **3** c **4** $\mu = \frac{11}{3}$

$\sigma = \sqrt{5}$

5 $\mu = -\frac{23}{24}$ $\sigma \simeq 2.51$

6 $\mu = \frac{5}{12}$ $\sigma \simeq 1.55$

7 **First:** a = 0.4

Second: $\mu = 3$ `$\sigma \simeq 1.55$

8 $\mu = 2.6$ $^2\sigma = 0.96$

9 **First:** $\frac{11}{16}$ **Second:** n = 2 , $\sigma^2 = 1$

10 **First:** $r = \frac{1}{3}$

Second: Probability distribution:

x_r	- 3	Zero	3	6
$f(x_r)$	$\frac{3}{9}$	$\frac{1}{9}$	$\frac{2}{9}$	$\frac{3}{9}$

$\mu = \frac{5}{3}$ $\sigma^2 \simeq 14.2$

11 a = 6 , $\sigma^2 = 1.55$

12 **First:** $C = \frac{13}{2}$

Second: $\mu = \frac{7}{3}$ $\sigma^2 = \frac{5}{9}$

13 **First :** C = 18

Second: Coefficient of variation = 37.7 %

14 **First:** m = - 1

Second: $\mu = \frac{5}{8}$, $\sigma^2 = \frac{135}{64}$

15 **First:** $C = \frac{20}{19}$

Second: $\mu = \frac{23}{19}$, $\sigma^2 = 1.2$

16 **First:** $L(x = 0) = \frac{1}{8}$ $L(x = 2) = \frac{5}{8}$

Second: Coefficient of variation = 132.3٪

17 **First:** $a = \frac{1}{12}$ ` k = 3 **Second:** $\sigma = 1.15$

Answers of exercises (3 - 3)

1 b **2** a **3** a

4 **First:** $\frac{1}{4}$ **Second:** $\frac{7}{2}$

5 **First:** $\frac{3}{4}$ **Second:** $\frac{5}{12}$

6 **First:** d (x) Density function **Second:** $\frac{20}{27}$

7 **First:** $\frac{4}{9}$ **Second:** $\frac{7}{9}$

8 **First:** $a = \frac{1}{8}$ **Second:** $\frac{1}{2}$

9 **First:** a = 0 **Second:** $\frac{1}{2}$

10 **First:** $a = \frac{1}{4}$ **Second:** $\frac{1}{2}$

11 **First:** k = 8 **Second:** $\frac{3}{16}$

12 **First:** $\frac{1}{4}$ **Second:** $\frac{7}{2}$

13 **First:** a = 1 **Second:** b = 2

Answers of the general exercises of unit 3

1 g **2** a **3** a

4

x_r	- 3	- 1	0	1	3
$f(x_r)$	$\frac{1}{9}$	$\frac{3}{18}$	$\frac{4}{9}$	$\frac{3}{18}$	$\frac{1}{9}$

5 $\therefore$ m = - 1

x_r	-2	-1	1	2
$f(x_r)$	$\frac{2}{16}$	$\frac{3}{16}$	$\frac{5}{16}$	$\frac{6}{16}$

6 $a = \frac{1}{7}$ $\frac{5}{14}$

(1)

x_r	1	2	3	4	5	6
$f(x_r)$	$\frac{1}{36}$	$\frac{3}{36}$	$\frac{5}{36}$	$\frac{7}{36}$	$\frac{9}{36}$	$\frac{11}{36}$

(2) **First:**

x_r	0	2	4
$f(x_r)$	$\frac{3}{9}$	$\frac{4}{9}$	$\frac{2}{9}$

Second: $\frac{7}{9}$

(3)

x_r	3	5	7	9	11
$f(x_r)$	$\frac{1}{9}$	$\frac{2}{9}$	$\frac{3}{9}$	$\frac{2}{9}$	$\frac{1}{9}$

$\mu = 7$, $\sigma \simeq 2.3094$

(4)

x_r	2	3	4	5	6	7	8
$f(x_r)$	$\frac{1}{16}$	$\frac{2}{16}$	$\frac{3}{16}$	$\frac{4}{16}$	$\frac{3}{16}$	$\frac{2}{16}$	$\frac{1}{16}$

$\mu = 5$, $\sigma \simeq 1.581$

(5) $b = \frac{1}{6}$, $a = 3$ $\sigma \simeq 1.08$

(6) **First:** $\frac{3}{8}$ **Second:** $\frac{27}{32}$

(7) **First:** $a = 2$ **Second:** $\frac{7}{16}$

Answers of the accumulative test of unit 3

(1) $\frac{5}{11}$ (2) 2.16 (3) 0.38 (4) 0.38

(5) **First:** $\frac{1}{3}$ **Second:** $\frac{2}{3}$

Unit 4: Normal distribution

Answers of exercises (4 - 1)

(1) a 0.3749 , 0.4922 b 0.0160 , 0.4484

c 0.516 , 0.901 d 0.9447 , 0.6971

e 0.0757 , 0.0669 f 0.1609 , 0.1337

g 0.0749 , 0.0202 h 0.8729 , 0.9898

i 0.7422 , 0.9222 j 0.3264 , 0.0548

(2) a 1.06 b -1.35 c 0.28 d 1.97

e - 0.85 f 1.32 g - 2.61

h - 0.84 i 0.65 j - 0.33

(3) a - 0.94 , 0.7818 b 2.04 , 0.2670

c 1.25 , 0.6524 d 1.83 , 0.9664

e 2.67 , 0.9154 f - 1.3 , 0.6766

(4) a $\sigma = 8$ b $\sigma = 7.5$ c $\mu = 40$

d $\mu = 60$ e $\mu = 50$ f 0.58

g k = 38 h k = 80 i k = 52

(5) a k = 137.5 b $\mu = 45$

c **First:** k = 10.5 **Second:** 0.8413

d 0.2902

e **First:** 1.91 **Second:** k = 1.65

f **First:** 0.1151 **Second:** 0.5403

g **First:** 0.0446 **Second:** 0.8185

h **First:** 0.8274 **Second:** 52.5

i **First:** 0.6247 **Second:** 0.8413

j **First:** 0.0401 **Second:** 0.6147

k **First:** 0.1056 **Second:** 10.5

Answers of exercises (4 - 2)

(1) 0.6247 , 841 families

(2) 49.67% , 159 students

(3) 62 students.

(4) 1547 students

(5) a 20 families b 465 L.E

(6) a 0.1056 b 179 families

(7) 0.9525 (8) 442 employees

(9) a 0.8413 b 0.0928

(10) a 0.6451 b 0.2119

c 0.2382

(11) a 977 young people

b 67 young people

(12) 16

1 (a) $\sigma = 11.36$ (b) 587 students
2 (a) 49.67 % (b) 159 students
3 $\sigma = 12.5$
4 5 students
5 971 cylinders
6 0.8417
7 $\mu = 170$
8 (a) 0.8413 (b) 0.9544
9 $\mu = 50$
10 (a) 0.5328 (b) 0.5987
11 (a) 0.9835 (b) 0.1949

Answers of the general exercises of unit 4

1 0.0228
2 $\mu = 77.6$
3 $\sigma = 7.5$
4 (a) 0.7734 (b) 0.8012 (c) 0.9332
5 $\mu = 16$, $\sigma = 4$
6 (a) 0.9544 (b) 0.1535
7 (a) $\mu = 25$ (b) $\mu = 20$
(c) $\sigma = 10$
8 (a) 0.0228 (b) K = 40
9 (a) a = 99.36 (b) 0.0062
(c) 0.0228 (d) 0.7888
10 (a) $\mu = 48$ (b) 0.4013
11 (a) 1.25 (b) 1.22
12 (a) 0.5 (b) 0.5
(c) 0.6826 (d) 0.0228
(e) 0.8185 (f) 0.3830
13 (a) a = 0.93 (b) a = 1.48
14 K = 90.6
15 (a) $\mu = 55$ (b) $\mu = 37.5$
16 0.3153
17 (a) 0.0228 (b) 89.44 %
(c) 187 employees
18 (a) 154 employees
(b) 20 employees
19 (a) 0.8413 (b) 95.44 %

Answer of the accumulative test

1 **First:** (a) $\frac{2}{3}$ (b) $\frac{1}{3}$
Second: Yes because their intersection is ϕ
2 (a) $\frac{1}{6}$ (b) $\frac{1}{2}$
3 (a) 0.8
(b) y = 0.0428571 x + 8.571435
(c) y = 94.29
4 (a) k = 1 (b) $\frac{1}{9}2$
(c) $\frac{4}{9}$
5 **First:** (a) $\frac{1}{6}$ (b) $\frac{2}{3}$
Second: $\frac{1}{12}$
6 170 cm
7 0.52857 direct